ESTABLISHING THE EFFICACY OF PSYCHOLOGICAL TREATMENTS AND UNDERSTANDING PSYCHOLOGICAL MECHANISMS UNDERPINNING EMOTIONAL DISTRESS IN BREAST CANCER

Thesis submitted in accordance with the requirements of the University of Liverpool

for the degree of Doctor of Philosophy

by James Temple

December 2018

Abstract

Previous meta-analyses of psychological treatments for emotional distress in breast cancer (BCa) conclude that efficacious psychological treatments exist. Consequently, their implementation in routine care is widely promoted by healthcare policy. However, decisions to implement and recommend these treatments should be based on high quality empirical data. It is widely recognised that high-quality randomised controlled trials (RCTs) provide the most reliable evidence of treatment efficacy. Therefore, this thesis began with a scoping review of the relevant published literature to identify potential systematic reviews on the quality of psychotherapy RCTs for emotional distress in BCa. Nine reviews were identified; all of which inadequately assessed the methodological quality of the included RCTs. A systematic review of the quality of psychotherapy RCTs of emotional distress in BCa was therefore conducted. 91 RCTs were eligible. Overall, methodological quality was low. Numerous methodological limitations were identified; three limitations were of considerable concern. Only 51% of RCTs used a treatment manual; only 15% specified inclusion criterion that participants were distressed; and only 11% reported the clinical significance of findings. If relevant health policies are to be adequately empirically informed, meta-analyses must account for these methodological limitations.

As previous meta-analyses have failed to account for important methodological limitations; none focused specifically on clinically distressed patients, excluded non-manualised treatments, or examined whether treatment effects were clinically significant, their practical relevance is questionable. Therefore, an individual patient data meta-analysis of RCTs of manualised psychological treatments for emotional distress in BCa patients was conducted. Treatment efficacy was evaluated using both effect size and clinical significance analyses; and analyses were conducted on the total sample (including distressed and non-distressed patients), and the clinically distressed sub-sample. Individual patient data was collected for 17 (n=2,996) of 26 (n=5,049) eligible trials. In the total sample, controlled effect sizes comparing treated and control patients were non-significant. In the clinical significance analysis, statistically significant benefits at post-treatment belied small differences and no differences remained at follow-up. In the distressed sub-sample, controlled effect sizes favouring treated patients were significant at post-treatment, but not at follow-up. In the clinical significance analysis, overall recovery was low: at post-treatment only 28-32% of

treated patients recovered compared to 17-27% of controls. At follow-up, only 24-33% of treated patients recovered compared to 24-34% of controls. Psychological treatments do not appear to alleviate emotional distress for most BCa patients. This contradicts previous meta-analyses and highlights the pressing need to develop more efficacious psychological treatments for distressed BCa patients.

To develop more efficacious psychological treatments for distressed BCa patients, a better understanding of the psychological processes associated with the development and maintenance of emotional distress in BCa is needed. Two psychological constructs that could explain distress in BCa survivors are beliefs about thinking, known as metacognitive beliefs; and the tendency to find uncertain situations distressing, known as intolerance of uncertainty. Therefore, an experience sampling methodology study examined the role of metacognitive beliefs and intolerance of uncertainty in predicting emotional distress and repetitive negative thinking (i.e. worry & rumination) in BCa survivors. Neither metacognitive beliefs nor intolerance of uncertainty predicted emotional distress. However, negative metacognitive beliefs about the uncontrollability and danger of repetitive negative thinking predicted repetitive negative thinking after controlling for intolerance of uncertainty and were a better predictor than baseline repetitive negative thinking, age at diagnosis, tumour stage at diagnosis, employment status, living alone or not, and time since finishing adjuvant therapy.

Psychological treatments of greater efficacy are urgently needed for BCa patients with emotional distress. Improvements in understanding the psychological processes underpinning emotional distress in BCa may lead to more efficacious treatment development.

Table of Contents

List of Figures

List of Tables

List of Appendices

Glossary

Abbreviation	Definition
AIC	Akaike's Information Criterion
AYA	Adolescent and young adult
BCa	Breast cancer
BDI	Beck Depression Inventory
CAS	Cognitive attentional syndrome
CAS-1	Cognitive Attentional Syndrome Scale
CBT	Cognitive behavioural therapy
CES-D	Centre for Epidemiological Studies Depression Scale
CI	Confidence interval
EMA	Ecological momentary assessment
EN	Edwards-Nunnally
ESM	Experience sampling methodology
FCR	Fear of cancer recurrence
GAD	Generalised anxiety disorder
GLN	Gullisken-Lord-Novick
HA	Hageman and Arrindell
HADS	Hospital Anxiety and Depression scale
HADS-A	Hospital Anxiety and Depression scale – anxiety subscale
HADS-D	Hospital Anxiety and Depression scale – depression subscale
HADS-T	Hospital Anxiety and Depression scale total
HLM	Hierarchical linear modelling
HAM-D	Hamilton Rating Scale for Depression
ICC	Intra-class correlation coefficient
IPD	Individual patient data
IPD-MA	Individual patient data meta-analysis
ITT	Intention to treat
IU	Intolerance of uncertainty
IUS	Intolerance of Uncertainty Scale
MBT	Mindfulness based therapy
MCQ-30	Metacognitions questionnaire-30
MeSH	Medical Subject Headings

Abbreviation	Definition
MRC	Medical Research Council
NCCN	National Comprehensive Cancer Network
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NK	Nunnally-Kotsch
OCD	Obsessive compulsive disorder
PD	Parkinson's disease
POMRF	Psychotherapy Outcome Study Methodology Rating Form
POMS	Profile of Mood States
POMS-TMD	Profile of Mood States Total Mood Disturbance
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSWQ	Penn State Worry Questionnaire
PTSD	Post-traumatic stress disorder
QoL	Quality of life
RCI	Reliable change index
RCT	Randomised controlled trial
RCT-PQRS	Randomized Controlled Trial Psychotherapy Quality Rating Scale
RD	Risk difference
RoB	Risk of bias
RRS	Ruminative response scale
SD	Standard difference
SE	Standard error
SMD	Standardised mean difference
S-REF	Self-regulatory executive functioning

Acknowledgements

I would like to thank the following people for helping me complete my thesis:

Thank you to my supervisors, Professor Catrin Tudur-Smith, Dr Peter Fisher and Professor Peter Salmon who have supported and helped me throughout my PhD. I am especially thankful to Peter Fisher and Peter Salmon; both have encouraged and challenged me, provided constructive criticism where necessary, and assisted in my professional development. I wouldn't have been able to complete this thesis without their help. Peter Salmon, your advice and attention to detail has helped me become much more appreciative of research. Peter Fisher, your thought-provoking questions and kind words of wisdom have helped me become a much more critical thinker.

I would also like to thank the following people:

- Mr Chris Huntley and Dr Angela Byrne who assisted with screening and data extraction.
- Dr Gemma Cherry who offered a lot of support, proof-read my work and provided feedback whenever I asked.
- Dr Maria Sudell for assisting with the ESM analysis.
- Ms Laura Hope-Stone and Ms Lucy Brading for sharing an office and putting up with me throughout the years and offering advice.
- The breast cancer nurses at the Royal, especially Mrs Sue Holcombe for helping with recruitment in the ESM study.
- My parents for their practical and financial support over the years.

Lastly, I would like to thank my partner, Taegan Jones. Thanks for putting up with me (nothing to do with the PhD just in general). Also, thanks for putting up with my absence when I've worked late and on weekends. Although this didn't seem to bother you as much as I expected! Thanks for all your support throughout!

Published and Submitted Publications from Thesis

Temple, J., Salmon, P., Tudur-Smith, C., Huntley, C. D., & Fisher, P. L. (2018). A systematic review of the quality of randomized controlled trials of psychological treatments for emotional distress in breast cancer. *Journal of Psychosomatic Research, 108*, 22-31. https://doi.org/10.1016/j.jpsychores.2018.02.013

Temple, J., Salmon, P., Tudur-Smith, C., Huntley, C.D., Byrne, A., & Fisher, P (submitted) Manualized psychological treatments do not alleviate emotional distress in most breast cancer patients: An individual patient data meta-analysis.

Declaration

I, James Temple, declare that I am the author of this thesis, that, unless otherwise stated, all references cited have been consulted by me, that, unless otherwise stated, the work of which this thesis is a record has been done by myself and has not been previously accepted for a higher degree.

James Temple

December 2018

Thesis Overview

The thesis reports four studies conducted to meet two broad aims: i) advance knowledge and understanding regarding the efficacy of psychological treatments for emotional distress in breast cancer (BCa); and ii) identify the psychological processes associated with the development and maintenance of emotional distress in BCa survivors.

Chapter 1 provides a general overview of the impact of emotional distress among BCa patients. Screening methods for emotional distress in BCa; psychological treatments currently available to BCa patients to alleviate emotional distress; and conclusions of previous meta-analyses regarding the efficacy of psychological treatments for emotional distress in BCa are discussed. The chapter ends by highlighting the impact that methodological quality of RCTs included in meta-analyses has on their findings and subsequent conclusions.

Chapters 2 to 5 address the first aim of this thesis by challenging the widespread assumption that efficacious psychological treatments for emotional distress in BCa exist. Chapter 2 presents a scoping review of systematic reviews evaluating the quality of psychotherapy RCTs for emotional distress in BCa. Chapter 3 presents a systematic review of the quality of RCTs of psychological treatments for emotional distress in BCa. Following a review and critique of the Jacobson clinical significance method in chapter 4, chapter 5 presents an individual patient data meta-analysis (IPD-MA) of the efficacy of manualised psychological treatments for emotional distress in BCa patients.

Chapters 6 to 8 address the second aim of this thesis by attempting to identify the psychological processes associated with the development and maintenance of emotional distress in BCa survivors. Chapter 6 provides an overview of two psychological models that could account for the development and maintenance of emotional distress in BCa: the intolerance of uncertainty model (Dugas, Gagnon, Ladouceur, & Freeston, 1998) and the Self-Regulatory Executive Function (S-REF) model (Wells & Matthews, 1994, 1996; Wells, 2009). Following a brief overview of Experience sampling methodology (ESM) in chapter 7, chapter 8 presents an ESM study examining the key constructs of the S-REF model (metacognitive beliefs) and the intolerance of uncertainty model (intolerance of uncertainty) in predicting

repetitive negative thinking (i.e. worry and rumination) and emotional distress in BCa survivors.

Finally, chapter 9 summarises the overall findings of the thesis. The limited efficacy of available psychological treatments and the implications this has on future research and practice are discussed.

Chapter 1. An Overview of Breast Cancer and Emotional Distress

1.1 What is breast cancer?

Cancer is a term used for a collection of diseases in which abnormal cells grow in an uncontrollable manner and invade healthy cells in the body. Breast cancer (BCa) is an uncontrolled growth of cells within the breast. BCa usually begins in the cells of the ducts, the passages that drain milk from the lobules to the nipple (Ductal Carcinoma); but can also begin in the cells of the lobules, the glands that produce breast milk (Lobular Carcinoma). If left untreated, primary BCa cells (non-metastatic BCa) can spread, or metastasise, through the lymphatic or blood system to other body parts (metastatic BCa).

1.2 Breast cancer prevalence and survival

Although both genders can develop BCa, over 99% of cases are in women (Siegel, Miller, & Jemal, 2017). Globally, BCa is the most common cancer amongst women, accounting for almost 1 in 4 female cancer cases (Bray et al., 2018). An estimated 2.1 million women are diagnosed with BCa worldwide each year (Bray et al., 2018) whilst one in eight women in the United States and UK are diagnosed with BCa in their lifetime (Siegel, Miller, & Jemal, 2018; Cancer Research UK, 2015). In the early 20th century, the life-expectancy after BCa diagnosis was approximately three years (Charache, 1932). However, improvements in detection methods and advances in medical treatment over the past century have significantly increased survival rates in BCa (Page & Adler, 2008). It is estimated that there are around 3.5 million BCa survivors in the United States (Miller et al., 2016) and 570,000 in the UK (Maddams et al., 2009); and the 5, 10 and 15-year survival rates for BCa are now 90%, 83% and 78%, respectively (Howlader, Noone, & Krapcho, 2015). This improvement in survival has led to an increased focus on understanding the psychosocial consequences of BCa.

1.3 Emotional distress in breast cancer

The term *'distress'* has many different meanings. The National Comprehensive Cancer Network (NCCN) Distress Management Panel defines distress as;

> *"…a multi-determined unpleasant emotional experience of a psychological (cognitive, behavioral, emotional), social, and/or spiritual nature that may interfere with the ability to cope effectively with cancer, its physical symptoms*

and its treatment. Distress extends along a continuum, ranging from common normal feelings of vulnerability, sadness and fears to problems that can become disabling, such as depression, anxiety, panic, social isolation, and spiritual crisis" (National Comprehensive Cancer Network, 2003).

However, prevalence of emotional distress in BCa is usually defined in terms of general distress, anxiety disorders, and/ or depressive disorders opposed to an all-encompassing definition as above.

Emotional distress is common throughout the BCa disease trajectory. Approximately 50% of BCa patients report clinical levels of emotional distress shortly after diagnosis, and 25% of patients do so in each of the second, third and fourth years after diagnosis (Bouchard et al., 2016; Burgess et al., 2005). In addition, life-time prevalence of BCa related post-traumatic stress disorder (PTSD) is 10-12% - almost twice that of the general female population (Andrykowski & Kangas, 2010; Frans, Rimmö, Åberg, & Fredrikson, 2005). Consequently, emotional distress is recognised globally as the sixth vital sign of BCa care used alongside pulse, temperature, respiration, blood pressure and pain (Bultz & Carlson, 2006) to assess patient's health and well-being.

1.4 Trajectory of emotional distress after breast cancer diagnosis

For most BCa patients, emotional distress shortly after diagnosis represents a normal and potentially adaptive stress response to a traumatic and threatening event, and would be expected to resolve spontaneously without specialist help (Brennan & Moynihan, 2004). In a study investigating trajectories of emotional distress in BCa across a five-year period, the most spontaneous reduction in distress prevalence occurred in the first 3 months following diagnosis (Helgeson, Snyder, & Seltman, 2004). However, distress does not always resolve naturally. Some patients remain distressed or become distressed at a later stage. Henselmans et al. (2010) assessed distress at five key points in the BCa disease trajectory (i.e. soon after diagnosis, post-surgery, immediately after adjuvant therapy, and two- and six-months after adjuvant therapy). Although most BCa patients experienced no distress (36%) or returned to normal within 6 months following completion of adjuvant therapy (33%), a small minority experienced persistent distress throughout the trajectory (15%) or only became distressed soon after adjuvant therapy ended (15%). Evidence of these distress trajectories has been supported by other studies in BCa (Boyes et al., 2013; Deshields,

Tibbs, Fan, & Taylor, 2006; Millar, Purushotham, McLatchie, George, & Murray, 2005). Thus, it cannot be assumed that, following a diagnosis of BCa, all patients follow one and the same trajectory of emotional distress.

1.5 Impact of emotional distress in breast cancer

One of the highest unmet needs raised by BCa patients relates to the psychological domain (Armes et al., 2009; Okediji, Salako, & Fatiregun, 2017; Sanson-Fisher et al., 2000). Almost half (45%) of BCa patients report that the emotional sequelae of cancer has the greatest negative impact on their quality of life (QoL; Macmillan Cancer Support, 2016).

Moreover, BCa patients experiencing emotional distress are less compliant with adjuvant therapy (Colleoni et al., 2000) and more likely to use community health or accident and emergency services (Carlson & Bultz, 2004), placing greater demands on health-care provision and costs. Emotional distress also increases the risk of suicide (Reddy, 2010), the incidence being 1.4 times higher in BCa patients who are at least one year following diagnosis than the general population (Misono, Weiss, Fann, Redman, & Yueh, 2008). Subsequently, healthcare policies worldwide specify that BCa patients should be screened for emotional distress at key points in the disease trajectory (i.e. immediately after diagnosis, transition into adjuvant therapy, and shortly after completion of adjuvant therapy) and have access to psychological treatment if required (Holland, Watson, & Dunn, 2011; Page & Adler, 2008; Tit et al., 2017).

1.6 Screening for emotional distress in breast cancer

Screening for emotional distress in BCa is heavily orientated around the diagnostic model of identifying need, according to which detecting emotional distress indicates the need for explicit psychological treatment (Salmon, Clark, McGrath, & Fisher, 2015). Many healthcare policies follow a tiered-model approach. In the UK, the National Institute for Health and Care Excellence (NICE) recommend a four-tier model (National Institute for Clinical Excellence, 2004). At level 1, compassionate communication and general psychological support is available to patients regardless of distress by front line staff (e.g. doctors, nurses and allied healthcare professionals). At level 2, health & social care professionals with additional expertise screen patients

for distress using a self-report measure, such as the distress thermometer (a single-item self-rated visual analogue scale), to rapidly identify patients reporting distress indicative of needing psychological treatment. If necessary, basic psychological techniques, such as problem-solving and psycho-education are provided. At level 3 and 4, patients with severe distress scores are referred to trained and accredited psychological professionals (i.e. counsellors, psychiatrists or health and clinical psychologists) for a psychological assessment and, if appropriate, offered specialist psychological treatment (e.g. cognitive behavioural therapy; CBT).

However, focusing solely on intensity of distress when screening is not the most comprehensive approach, and other factors such as desire for help (felt need), demand and enthusiasm for psychological services (expressed need), and identification of adaptive and maladaptive emotional responses also need to be considered (Dekker et al., 2017; Girgis, Smith, & Durcinoska, 2018; Salander, 2017; Salmon et al., 2015; Schaeffeler et al., 2015; Tondorf et al., 2018).

At what point in the disease trajectory patients are screened is also important. Screening and addressing emotional distress shortly after diagnosis is not always appropriate. Recently diagnosed BCa patients often view distress as a temporary and understandable reaction to cancer and do not want early psychological intervention (Baker et al., 2016). Interrupting this normative experience may disrupt the equilibrium of adaptive processing (Brennan & Moynihan, 2004; Dekker et al., 2017). Moreover, PTSD literature suggests that early intervention can do more harm than good (Litz, Gray, Bryant, & Adler, 2002; Rose, Bisson, & Wessely, 2003). Thus, while there is a consensus about the importance of identifying patients in need of psychological support, the most appropriate method of determining need has not been established.

1.7 Psychological treatments for emotional distress in breast cancer

In recognition of the persistent emotional sequelae of BCa, numerous psychological treatments for emotional distress in BCa patients have been evaluated. The most commonly evaluated approaches are cognitive behavioural therapy (CBT), psychoeducation, and supportive therapy (Fors et al., 2011). In recent years, several studies have also evaluated mindfulness-based therapy (MBT). The overwhelming consensus in psycho-oncology research is that these treatments are efficacious and

alleviate emotional distress in BCa. Two Cochrane reviews (Jassim, Whitford, Hickey, & Carter, 2015; Mustafa, Carson-Stevens, Gillespie, & Edwards, 2013) and eight additional meta-analyses (Cobeanu & David, 2018; Duijts, Faber, Oldenburg, van Beurden, & Aaronson, 2011; Naaman, Radwan, Fergusson, & Johnson, 2009; Tatrow & Montgomery, 2006; Xiao et al., 2017; Ye et al., 2018; Zimmermann, Heinrichs, & Baucom, 2007) of randomised controlled trials (RCTs) evaluating the efficacy of psychological treatments compared to control conditions (assessment only, wait-list controls, treatment as usual, or active controls) have been conducted, with all ten concluding that efficacious psychological treatments exist. Consequently, healthcare policies and clinical practice guidelines internationally specify that psychological treatment should be available to BCa patients as part of their routine care throughout the disease trajectory (Dauchy et al., 2012; Holland et al., 2011; Howell et al., 2009; Li et al., 2016; National Breast Cancer Centre, 2003; National Comprehensive Cancer Network, 2003; National Institute for Clinical Excellence, 2004; Page & Adler, 2008; Reese, Weis, Schmucker, & Mittag, 2017; Tit et al., 2017). However, decisions to implement and recommend these treatments should be based on high quality empirical data.

1.8 Importance of methodological quality of randomised controlled trials in breast cancer

Confidence in conclusions of previous meta-analyses relies on the quality of their evidence-base. It is widely recognised that high-quality RCTs provide the most reliable evidence of treatment efficacy; and that low-quality RCTs tend to overestimate treatment effects (Barth et al., 2016; Bohlmeijer, Prenger, Taal, & Cuijpers, 2010; Bolier et al., 2013; Cuijpers, van Straten, Bohlmeijer, Hollon, & Andersson, 2010; Huhn et al., 2014; Klein, Jacobs, & Reinecke, 2007). A recent meta-analysis of 115 RCTs evaluating the efficacy of psychological treatments for depression found that effect sizes were significantly smaller in high-quality RCTs (d=0.22) than low-quality RCTs (d=0.74; Cuijpers et al., 2010). Treatment effects can also be overestimated even when RCTs neglect individual aspects of quality. For example, larger effects sizes have been found in RCTs not using blinded outcome assessors (Jauhar et al., 2014), not using intention to treat (ITT) analyses (Frühauf, Gerger, Schmidt, Munder, & Barth, 2013; Klein et al., 2007), and not adequately randomising participants to conditions (Gellatly et al., 2007).

The overestimation of treatment effects is not the only concern of poor quality RCTs. Poor quality also undermines the confidence in the conclusions that can be drawn from RCTs (Gerber et al., 2011; Öst, Havnen, Hansen, & Kvale, 2015; Thoma et al., 2012). For example, if concomitant treatments (i.e. pharmacotherapy or additional psychotherapy) are not controlled for, it is difficult to determine the impact of the specific treatment being assessed; if the control condition receives assessment only, it remains unclear whether observed effects are due to specific treatment ingredients or features that are common across different therapies; and if the sample is not representative of those seeking treatment in practice, researchers cannot be confident that the findings are generalisable to the clinical setting.

It is therefore crucial that the quality of RCTs of psychological treatments is known if policymakers and clinicians are to make informed decisions about the implementation of, and referral to, psychological treatments in clinical services. Assessing the methodological quality of RCTs has been fundamental to advancing the scientific credibility and reporting standards of psychotherapy outcome trials in mental health settings (Cuijpers et al., 2010; Gerber et al., 2011; Öst et al., 2015; Thoma et al., 2012). For example, it appears that as the quality of psychotherapy RCTs for depression have improved, the magnitude of treatment effects have diminished.

It is clear that the methodological quality of RCTs included in meta-analyses can have a substantial impact on their subsequent conclusions. Therefore, the general conclusion that efficacious psychological treatment for emotional distress in BCa exist can only be ascertained if the RCTs that these conclusions are based on are of good quality.

1.9 Summary

Emotional distress in BCa is a significant problem worldwide. Therefore, access to efficacious psychological treatments for emotional distress in BCa is essential. Meta-analyses of RCTs conclude that efficacious psychological treatments for emotional distress in BCa exist. However, confidence in these conclusions is determined by the quality of their evidence-base. Thus, the quality of the evidence-base with which previous conclusions are based needs to be known in order to determine the level of confidence one can have in the general conclusion that efficacious psychological treatments for emotional distress in BCa exist.

Chapter 2. Study 1: A Scoping Review of Reviews Assessing the Quality of Randomised Controlled Trials of Psychological Treatments for Emotional Distress in Breast Cancer

2.1 Introduction

To determine the quality of the evidence-base with which previous meta-analytic conclusions in BCa are based, a scoping review of the relevant published literature was conducted to identify potential systematic reviews on the quality of psychotherapy RCTs for emotional distress in BCa. This chapter presents the findings of this scoping review. First, however, an overview of the different methods used to assess RCT quality, and the importance of assessing specific design elements when evaluating RCT quality is presented.

The term 'methodological quality' is a multidimensional construct which encompasses design, implementation, analysis, reporting, replicability, and generalisability (Sundell & Åhsberg, 2016). Within the healthcare literature there are inconsistencies in how methodological quality is defined, and the terms 'quality' and 'bias' are often used interchangeably (Hartling et al., 2009). However, these terms should be distinguished from each other. Bias refers only to the internal validity of a trial (i.e. reliability or accuracy of trial results; Higgins et al., 2011); while quality refers to the extent to which a trial was conducted to the highest possible standards (Verhagen, de Vet, de Bie, Boers, & van den Brandt, 2001). Thus, the concept of methodological quality includes more than just internal validity. It also includes external validity (i.e. generalisability of trial results to the target population), construct validity (i.e. extent to which a trial measures the intended construct), and statistical conclusion validity (i.e. extent to which trial data can be regarded as revealing a link between independent and dependant variables; Berlin & Rennie, 1999; Farrington, 2003).

To assess trial quality, three types of assessment tools exist: quality scales, quality checklists, and domain-based evaluations (Jüni, Witschi, Bloch, & Egger, 1999; Zeng et al., 2015). Although often used interchangeably, these terms are distinct (Olivo et al., 2008). With quality scales, individual items are numerically rated and combined to give an overall quality score. With quality checklists, individual items are given a categorical rather than a numerical rating (i.e. yes or no/ reported or not reported). With domain-based evaluations, each domain (e.g. detection bias) is given a categorical rating opposed to each individual item within that domain. As neither quality checklists nor domain-based evaluations rate items numerically, generating an

overall quality score is not possible. However, some quality checklists and domain-based evaluations include a summary judgement enabling an overall assessment of trial quality. For example, the quality checklist developed by Beatty et al. (2018) categorises trials as high quality if at least 8 of the 11 quality criteria are met.

Conclusions about trial quality vary according to the design elements being assessed. For example, Armijo-Olivo et al. (2015) used two quality assessment tools to assess the quality of 97 physical therapy RCTs. The two assessment tools differed with regards to the design elements being assessed. Only 11 of the 97 RCTs that were of adequate quality according to one assessment tool were of adequate quality according to the other. Yet, there is extensive variation in the design elements assessed by different assessment tools. A recent analysis in health research found that 130 different design elements were assessed across 19 quality assessment tools (Armijo-Olivo, Fuentes, Ospina, Saltaji, & Hartling, 2013).

Consensus on the design elements of RCTs which determine methodological quality does not exist. However, it is broadly accepted that several design elements are essential; proper randomisation, clear description of the sample, use of power analysis, sample representative of the target population, use of valid and reliable and specific outcome measures, use of an adequate comparator condition, control of concomitant treatments, adequate length of follow-up, complete outcome data (i.e. handling of attrition), adequate statistical methods, and assessment of clinical significance (Gerber et al., 2011; Kocsis et al., 2010; Liebherz, Schmidt, & Rabung, 2016; Öst, 2008; Sundell & Åhsberg, 2016).

Assessment of the methodological quality of psychotherapy RCTs has to contend with several interacting components which can be difficult to operationalise (Munder & Barth, 2018), Thus, when conducting a psychotherapy RCT, items such as manualisation, therapist adherence and competence, number of therapists, and equality of therapy hours also need to be considered (Kocsis et al., 2010; Luborsky & DeRubeis, 1984; Öst, 2008; Perepletchikova, Treat, & Kazdin, 2007).

The following section provides an overview of the importance of design elements essential to the methodological quality of RCTs in general as well as those specific to psychotherapy RCTs.

2.1.1 Generic design elements

Proper randomisation

The aim of randomisation is to ensure that confounding variables are equally distributed between participants within each group, increasing the likelihood that any observed differences between groups are due to treatment (Suresh, 2011). Proper randomisation depends on two separate but interlinked features: generation of an unpredictable random sequence (i.e. random sequence generation); and concealment of the random sequence (i.e. allocation concealment). Generation of an unpredictable sequence ensures that any future assignments cannot be anticipated. Randomisation based on a single sequence of random assignments, known as simple randomisation, is the most reliable method (Higgins et al., 2011). However, when simple randomisation is used on small samples, sample sizes between groups may be imbalanced. Thus, methods such as blocked or stratified randomisation are required (Kang, Ragan, & Park, 2008). Concealment of the unpredictable sequence prohibits selective enrolment of participants. For example, it prevents participants assigned to the treatment group who may be viewed as 'inappropriate' from being excluded (Higgins et al., 2011). Central randomisation by an independent third party is the most reliable method of allocation concealment (Higgins et al., 2011).

Sample representative of the target population

To enhance the external validity of an RCT and ensure results are clinically meaningful, findings must be generalisable to patients who would be offered treatment in routine practice (Rothwell, 2005). One of the most frequently reported reasons for the underuse of guideline-recommended treatments is the lack of empirical evidence for their efficacy amongst the type of patients seen in practice (Garfield & Garfield, 2000). To increase generalisability of findings, recruitment for RCTs should occur across several sites and eligibility criteria should avoid exclusion of representative patients.

Clear description of sample

A comprehensive description of patients included in an RCT is needed to ensure accurate interpretation of results (Amundsen et al., 2018). If the sample is poorly described, it is difficult to determine the representativeness of the sample or compare

outcomes across RCTs. A complete description of the eligibility criteria and the characteristics of the recruited sample are essential.

Use of power analysis

Power of an RCT refers to the probability of detecting a significant difference when a significant difference exists. The stronger the power, the less likely the probability of type II error (i.e. a false negative result). The power of an RCT is directly proportional to the sample size; and should be determined a priori to be at least 80% (i.e. the sample is large enough to detect a significant difference that exists 80% of the time) by conducting data-informed power analysis (Biau, Kernéis, & Porcher, 2008). Post-hoc power analysis should be avoided as they are often incorrect and misleading (Goodman & Berlin, 1994; Levine & Ensom, 2001).

Valid, reliable and specific outcome measures

Outcome measures need to have good psychometric properties to avoid unreliable or invalid data (Souza, Alexandre, & Guirardello, 2017). Essential psychometric properties are reliability and validity (Lohr, 2002). Reliability refers to the extent to which repeated measurements produce consistent results. Two types of reliability are important for outcome measures: test-retest reliability (i.e. reproducibility of results over time) and interrater reliability (i.e. reproducibility of results across different raters). Validity refers to the extent to which an outcome measure measures what was intended. Three types of validity are important for outcome measures: construct validity (i.e. extent to which an outcome measure measures the intended construct), content validity (i.e. extent to which an outcome measure measures all elements of the intended construct) and criterion validity (i.e. extent to which scores on one outcome measure predicts those on another). Outcome measures also need to be specific (i.e. relevant to the population in question).

Use of an adequate comparator condition

One of the primary purposes of a comparator condition in an RCT is to control for threats to internal validity (Mohr et al., 2009). As estimates of treatment efficacy depend on the contrast being made between conditions, the choice of comparator is central to any conclusions drawn. The least stringent comparators are wait-list control (WLC) or assessment only conditions. Although these conditions may control for non-

specific factors such as instillation of hope and trust (Modi, Wagner, Smith, Kellermann, & Michaelis, 2017), they do not control for other non-specific factors such as therapeutic alliance. A more reliable comparator is standard care or 'treatment as usual' (TAU). TAUs are more common in psychotherapy RCTs opposed to pharmacological RCTs (Freedland, Mohr, Davidson, & Schwartz, 2011). Using a TAU addresses the question of whether a new treatment is more beneficial than the current one being implemented. However, TAU varies considerably across services, hospitals and countries (Dawson et al., 2009; Freedland et al., 2011; Mohr et al., 2009), and is often poorly defined (Smelt, van der Weele, Blom, Gussekloo, & Assendelft, 2010) making it difficult to interpret results. The most stringent comparator in pharmacological RCTs is a placebo control as they control for non-specific factors (Simmonds, 2010). However, as it is difficult to adequately blind patients to psychological treatment, it is difficult to translate placebo controls into psychotherapy RCTs. Therefore, the most stringent comparator in psychotherapy RCTs is either an active control (e.g. supportive therapy or nondirective therapy) or another bona-fide psychological treatment; both of which help ensure observed differences are due to specific opposed to non-specific treatment factors.

Control of concomitant treatment

If receipt of concomitant treatment (i.e. additional pharmacological or psychological treatment) is not controlled for, it is difficult to determine whether observed effects are due to the treatment being evaluated or additional treatments (Öst et al., 2015). One way to increase internal validity of an RCT is to exclude patients receiving concomitant treatments. Another way is to ensure that, if patients are receiving concomitant pharmacological treatment, the dosage is kept stable before and during the RCT. If concomitant treatments are permitted, it is imperative that these treatments are reported for each group to ensure accurate interpretation of results.

Adequate length of follow-up

To check whether any observed effects are sustained, outcomes need to be measured beyond immediately post-treatment. The longer the follow-up period, the better the indication of the sustainability of effects. Any treatments patients receive during the follow-up period (i.e. pharmacological or psychological) need to be reported to improve interpretation of results and increase internal validity.

Complete outcome data (i.e. handling of attrition)

Incomplete outcome data due to attrition increases the likelihood that observed effects are biased (Higgins et al., 2011). Patients often drop out of a specific group due to lack of treatment success (Klein, Stone, Hicks, & Pritchard, 2003; Liebherz et al., 2016; Swift, Callahan, & Levine, 2009). If these patients are excluded from analyses, conclusions may be inaccurate. Patients who drop-out may also differ from completers on specific sample characteristics. Therefore, analyses should be conducted on intent to treat (ITT) basis and sample characteristics of those who completed treatment should be compared to those who dropped out.

Adequate statistical methods

Statistical analyses that appropriately test the hypotheses of an RCT are needed to guarantee reliable results (Norström, 2015).

Assessment of clinical significance

While statistical significance and effect sizes provide valuable group information, they provide no information about individual variability in treatment response (Loerinc et al., 2015). This makes it difficult to interpret the practical value of findings. Therefore, an evaluation of clinical significance is needed to indicate the proportion of patients who benefit from treatment (see chapter 4 for more details).

2.1.2 Psychotherapy-specific design elements

Manualisation

Psychotherapy outcome research is uninformative if there is a lack of clarity over the nature of the treatment evaluated (Chambless & Hollon, 1998). Therefore, treatment manuals are crucial to standardising psychological treatment and allowing discrimination between alternatives. Treatment manuals improve dissemination and the quality of treatment delivery (Addis, Cardemil, Duncan, & Miller, 2006) and help bridge the gap between research and practice. In the United States, psychological treatments must be manualised to be considered "evidence-based" and in turn recommended for use in clinical practice (Chambless & Hollon, 1998).

Therapist training

A manual alone cannot ensure that a treatment is delivered as intended. Therefore, treatment must be delivered by therapists trained to provide psychological treatments (e.g. licensed clinical or health psychologists). Therapists should also be trained in the specific treatment being evaluated. For example, a clinical psychologist delivering MBT must also be specifically trained to deliver MBT.

Therapist adherence & competence

Well-trained therapists may still fail to deliver a treatment correctly. Thus, adherence to the treatment manual and therapy competence also need to be assessed. Adherence refers to the extent to which therapists remain faithful to the prescribed treatment procedures, whereas competence refers to the degree of skill and judgment that therapists display when delivering the treatment (Barber, Sharpless, Klostermann, & McCarthy, 2007).

Number of therapists

Accumulating evidence indicates that therapists have a significant effect on outcome, accounting for around 5–10 % of unexplained variance (Crits-Christoph et al., 1991; Kim, Wampold, & Bolt, 2006). Therefore, to avoid confounding between therapist and therapy method, treatments must be delivered by multiple therapists and therapists must be included as a random design factor in analysis.

Equality of therapy hours

Internal validity of an RCT can be compromised if the duration and intensity of treatment conditions are not matched. Merely spending time with a therapist can lead to significant improvements in treatment outcome (Freedland et al., 2011). Therefore, if a comparator condition receives markedly less therapy hours, it is difficult to judge whether observed effects are due to the specific treatment being investigated or additional time spent with a therapist. However, equality of therapy hours is not always feasible; for example, RCTs evaluating different doses of the same treatment.

2.1.3 Aims

To be confident in conclusions of psychotherapy RCTs, and reviews based on those RCTs, the above design elements need to be correctly implemented. Therefore, the aims of this scoping review were to identify any reviews assessing the methodological quality of psychotherapy RCTs for emotional distress in BCa, and evaluate the nature and quality of the assessment tools used to determine RCT quality in those reviews. As bias represents one aspect of quality, reviews assessing the risk of bias (RoB) were also included.

2.2 Method

2.2.1 Search strategy

Relevant studies were identified by systematic searches of the following electronic databases: PubMed, PsycINFO, Web of Science, Scopus, and the Cochrane Database of Systematic Reviews. Databases were searched from their inception until June 2018 using Medical Subject Headings (MeSH) terms where possible. The search used a combination of terms associated with reviews (including: *"review* OR meta**), BCa (including: "*breast neoplasms" OR "breast cancer"),* emotional distress (including: "*Depression"[Mesh] OR "depressive disorder"[Mesh] OR "depressive disorder" OR "anxiety"[Mesh] Or "anxiety disorders"[Mesh] OR "anxiety disorders" OR "anxiety" OR "depression" OR "emotional distress" OR "psychological distress"*), and psychological treatments (including: *"Psychotherapy"[Mesh] OR "psychotherapy" OR "psychological therapy" OR "counselling" OR "counselling" OR "psychological intervention" OR "cognitive behavioural therapy" OR "group therapy" OR "psychosocial therapy"*). No date restriction was applied but only English language articles were included. The final search strategy used for each database is available in Appendix 1.

2.2.2 Eligibility criteria

Review type: Reviews of RCTs published in English in a peer-reviewed journal. Reviews not focusing specifically on RCTs were excluded.

Participants: Reviews exclusively comprising adults aged 18 years or older with a histologically confirmed diagnosis of BCa.

Treatment: RCTs included in reviews evaluated a psychological treatment in at least one condition.

Control: RCTs included in reviews had a no treatment (usual care) control, an active (attention placebo) control, or an alternative treatment in at least one condition.

Outcomes: The primary and/or secondary outcome of the RCTs included in reviews was emotional distress, defined as anxiety, depression, general mood, or global emotional distress.

Quality assessment: The methodological quality of RCTs included in reviews was assessed using a quality or RoB scale, checklist, or domain-based evaluation.

2.2.3 Review selection

After removing duplicates, a single reviewer (JT) screened titles and abstracts to remove irrelevant studies. Next, full-text of all potentially relevant papers was retrieved and assessed for inclusion by the same reviewer (JT). Uncertainties were discussed with a second reviewer (PF).

2.2.4 Data extraction

Data were extracted independently by one reviewer (JT) for all included reviews using a specially devised data extraction protocol. Data extracted was year of publication; number of RCTs included in the review; primary aim of review; outcomes measured in the review; type of quality assessment tool used; and items included in the quality assessment tool.

2.3 Results

Figure 2.1 presents an overview of the study selection procedure. The systematic search retrieved 864 citations of which 47 were potentially eligible following the title and abstract screening stage. After reading the full text, nine reviews published from 2009 to March 2018 were eligible and included. Six of the nine reviews included RCTs evaluating emotional distress whilst other RCTs within the same review measured different constructs such as QoL or survival. In this case, only the RCTs evaluating emotional distress were included in this review.

Characteristics of each review are presented in Table 2.1. Six reviews exclusively included; metastatic BCa patients (n=2), non-metastatic BCa patients (n=1), BCa patients who had completed surgery (n=2), or BCa patients who had completed adjuvant therapy (n=1). The remaining three reviews included BCa patients irrespective of disease stage or trajectory phase. Seven of the nine reviews assessed RoB only; and two assessed the broader domain of quality. To assess RoB, three assessment tools were used - one was a domain-based evaluation (Cochrane RoB tool; Higgins & Green, 2011) and two were scales (Jadad RoB scale; Fors' RoB scale; Fors et al., 2011; Jadad et al., 1996). To assess quality, two assessment tools were used – one was a checklist (Beatty's quality checklist; Beatty et al., 2018) and one was a scale (Naaman's quality scale; Naaman et al., 2009).

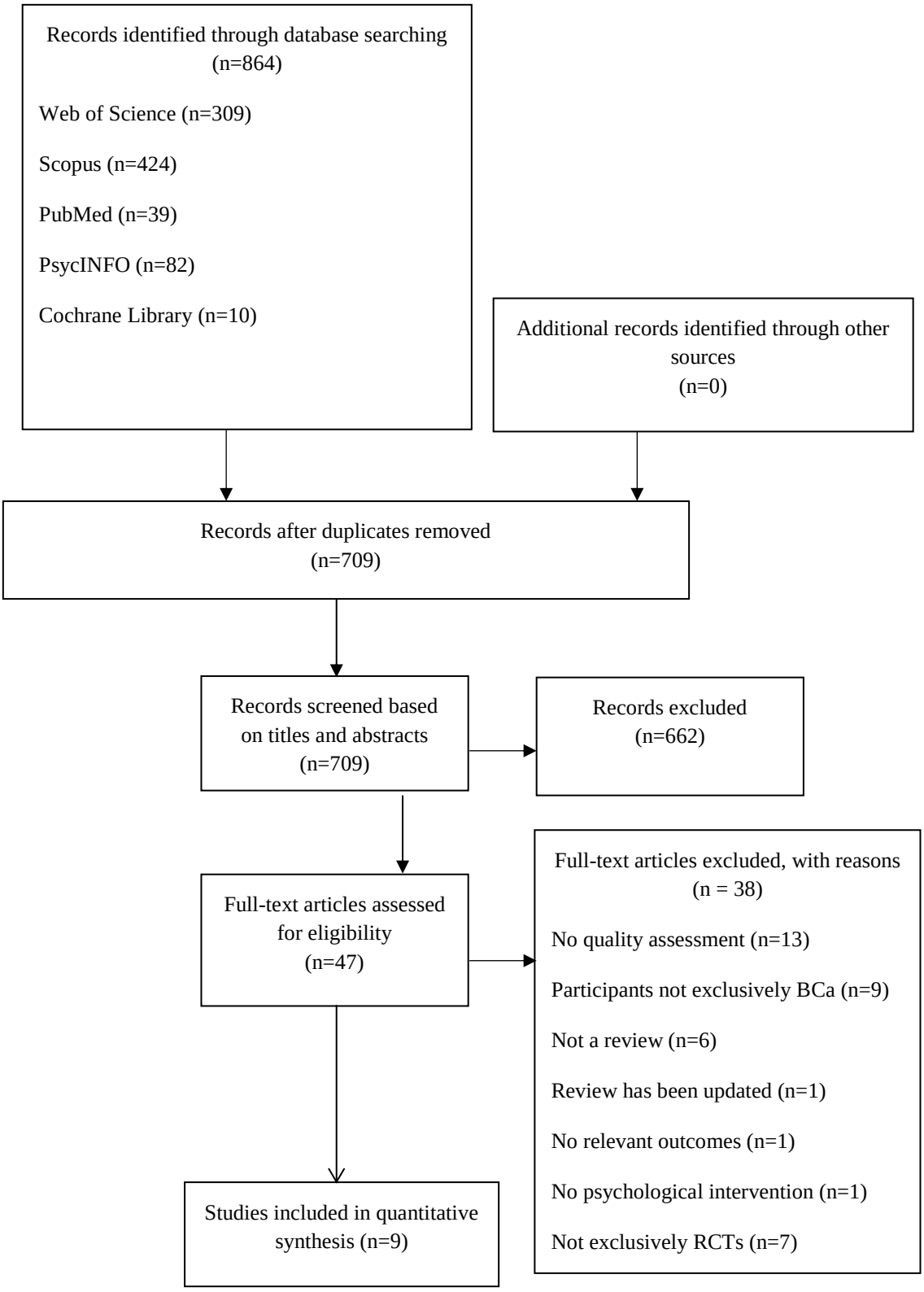


Figure 2.1: Flow chart showing review identification and selection

Table 2.1: Characteristics of included reviews

Review	Type of review	Objectives	Sample	Type of treatments	Outcomes measured	RCTs included	RCTs measuring distress in BCa	Year range for RCTs included
Beatty et al. 2018	Systematic review	To summarise the evidence-base of psychological treatments for women with metastatic BCa, by mode of delivery (group, individual, or low-intensity interventions)	Women with metastatic BCa	Any psychological treatment	**Primary outcomes:** Emotional distress, coping, QoL, survival	15	13	1981-2012
Fors et al. 2017	Systematic review	Determine the effectiveness of psychoeducation, CBT and social support interventions used in the rehabilitation of BCa patients	Female BCa patients who had completed surgery and adjuvant therapy	Psychoeducation, CBT and social support	**Primary outcomes:** QoL Mood *Depression, anxiety, stress* Clinical symptoms *Nausea, dizziness, health coping, work disability, leisure* Social functioning	18	17	1999-2008
Haller et al. 2017	Systematic review & meta-analysis	Systematically update the evidence for MBT in women with BCa	Women with BCa (Stage 0-IV) regardless of current treatment status	MBT	**Primary outcomes:** HRQoL, cancer-specific QoL, **Secondary outcomes:** Fatigue, sleep, stress, anxiety, depression	10	9	2009-2016
Jassim et al. 2015	Cochrane review	To assess the effects of psychological treatments on psychological morbidities, QoL and survival among women with non-metastatic BCa	Women with non-metastatic BCa	Any psychological treatment	**Primary outcomes:** Anxiety, depression, stress, mood disturbance **Secondary outcomes:** QoL, coping, adjustment, survival	28	28	1996-2013
Mustafa et al. 2015	Cochrane review	To assess the effects of psychological treatments on psychosocial and survival outcomes for women with metastatic BCa	Women with metastatic BCa	Any psychological treatment	**Primary outcomes:** Psychological outcomes	10	8	1989-2010

Review	Type of review	Objectives	Sample	Type of treatments	Outcomes measured	RCTs included	RCTs measuring distress in BCa	Year range for RCTs included
					Anxiety, depression, emotional distress, QoL, pain, condition-specific outcomes, relationship and social support Survival outcomes			
Naaman et al. 2009	Meta-analysis	Determine the overall efficacy of psychological treatments in BCa patients, specifically looking at three outcome variables: anxiety, depression, and QoL	Female BCa patients who had completed surgery	Psychological/behavioural treatments	**Primary outcomes:** Anxiety, depression, QoL	18	16	1981-2003
Xiao et al. 2017	Meta-analysis	Assess the efficacy of individually delivered CBT on improving the depressive symptoms of women with BCa	BCa patients who had completed surgery (time-frame not given)	CBT	**Primary outcomes:** Depression	10	10	1996-2015
Ye et al. 2018	Meta-analysis	Examine the effect of CBT on QoL and psychological health of BCa patients	Female BCa patients and survivors (survivors not defined)	CBT	**Primary outcomes:** QOL, depression, anxiety, stress	10	10	2003-2015
Zhang et al. 2016	Systematic review & meta-analysis	Quantify the effects of MBT on physical health, psychological health and Qol in patients with BCa	Women with BCa (Stage 0-IV) regardless of current treatment status	MBT	**Primary outcomes:** Psychological health *Anxiety, depression, spirituality, emotional well-being, stress, fear of recurrence* QoL Physiological health	7	5	2009-2014

Note. CBT = cognitive behavioural therapy; QoL = quality of life; HRQoL = health related quality of life; MBT = mindfulness-based therapy; BCa = breast cancer; RCT = randomised controlled trial

2.3.1 Quality assessment tools and reviews

The quality findings from each review are presented in Table 2.2. A summary of each assessment tool followed by a summary of the findings from each review using that assessment tool is provided below.

2.3.1.1 Cochrane Risk of Bias tool

The Cochrane RoB tool (Higgins & Green, 2011) assesses seven domains: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting (i.e. reporting only a sub-set of outcomes, depending on the nature and direction of the results), and a supplementary domain 'other bias'. For each RCT, users assign a judgement of 'high', 'low' or 'unclear' RoB to each domain. The guidelines recommend that an RCT has a 'low RoB' if a low RoB is found for all key domains, an 'unclear RoB' if an unclear RoB is found for one or more key domains, and a 'high RoB' if a high RoB is found for one or more key domains (Higgins & Green, 2011).

2.3.1.2 Reviews using the Cochrane Risk of bias tool

Mustafa et al., 2013

Mustafa et al. (2013) conducted a Cochrane review evaluating ten RCTs of psychological treatments for emotional distress and survival in metastatic BCa patients. Eight of the ten RCTs included emotional distress as an outcome variable. The RoB was unclear in all eight RCTs. Five of the seven domains (random sequence generation, allocation concealment, blinding of participants and personnel, incomplete outcome data, and selective reporting) were rated as having a low RoB in most RCTs; one domain (blinding of outcome assessment) was rated as having an unclear RoB in most RCTs; and one domain ('other biases') received a mixture of low and unclear RoB ratings across RCTs.

Jassim et al, 2015

In another Cochrane review, Jassim et al. (2015) evaluated 28 RCTs of psychological treatments for emotional distress in non-metastatic BCa patients. Nineteen RCTs had an unclear RoB and nine had a high RoB. Five of the seven domains (random sequence generation, allocation concealment, blinding of participants and personnel, blinding of

outcome assessment, and incomplete outcome data) were rated as having an unclear RoB in most RCTs, and two (selective reporting and 'other biases') were rated as having a low RoB in most RCTs.

Haller et al., 2017

Haller et al. (2017) carried out a meta-analysis of 10 RCTs of MBT for health-related QoL in BCa. Secondary outcomes included fatigue, sleep, stress, safety, and emotional distress. Nine of the ten RCTs included emotional distress as an outcome variable. All nine RCTs had a high RoB. Despite this, three of the seven domains (random sequence generation, incomplete outcome data, and 'other biases') were rated as having a low RoB in most RCTs; three (allocation concealment, blinding of participants and personnel, and blinding of outcome assessment) were rated as having an unclear RoB in most RCTs; and one (selective outcome reporting) was rated as having a high RoB in most RCTs.

Ye et al., 2018

Ye et al. (2018) conducted a meta-analysis of 10 RCTs of CBT for emotional distress in BCa. Eight RCTs had a high RoB, one had a low RoB, and one was rated as having an unclear RoB. Although most RCTs had a high RoB, five of the seven domains (random sequence generation, blinding of participants and personnel, incomplete outcome data, selective outcome reporting, and 'other biases') were rated as having a low RoB in most RCTs, while two (allocation concealment, blinding of outcome assessment) received mixed ratings across RCTs.

2.3.1.3 The Jadad scale

The Jadad scale (sometimes referred to as the Oxford quality scoring system; Jadad et al., 1996) is the most frequently used RoB scale in healthcare research (Olivo et al., 2008). It consists of five questions relating to three bias domains; random sequence generation, blinding of participants and personnel, and incomplete outcome data. Each question is answered 'yes' or 'no'. Each 'yes' scores one point and each 'no' scores zero points, producing a total score ranging from 0-5. A cut-off score of ≥ 3 out of 5 is recommended for an RCT to be classed as adequate quality (Jadad et al., 1996). The reliability of the Jadad scale has been disputed with interrater reliability ranging from low to high (Kappa:0.37-0.89) (Clark et al., 1999; Latronico et al., 2002).

2.3.1.4 Reviews using the Jadad scale

Zhang et al., 2017

Zhang, Xu, Wang, and Wang (2016) conducted a meta-analysis of seven RCTs of MBT for physical health, psychological health and QOL in BCa. Five of the seven RCTs included emotional distress as an outcome variable. Zhang et al. (2016) defined high-quality RCTs as those with a score of ≥4. According to this cut-off, two RCTs were high-quality; and according to the pre-defined cut-off score of ≥ 3, three RCTs (including the 2 high-quality RCTs) were of adequate quality. Four of the five RCTs failed to perform ITT analysis; three failed to ensure the outcome assessor was blinded, specify the method of randomisation, or ensure allocation concealment; and one failed to report the number of withdrawals or dropouts.

Xiao et al., 2017

Xiao et al. (2017) conducted a meta-analysis of 13 RCTs of CBT for depression in BCa patients after surgery. According to the pre-defined cut-off score of ≥3, all 13 RCTs were of adequate quality. No further information on trial quality was provided.

2.3.1.5 Fors' risk of bias scale

Fors et al. (2011) devised their own assessment scale to assess bias. It was developed based on the Cochrane RoB tool (Higgins & Green, 2011) and consists of 11 questions relating to seven bias domains (random sequence generation; allocation concealment; blinding of participants and personnel; blinding of outcome assessment; incomplete outcome data; statistical methods; and reliable outcome measures). Each question is answered 'yes', 'no', or 'unclear'. Each 'yes' receives one point and each 'no' or 'unclear' receives zero points, producing a total score ranging from 0-11. An RCT with a score of ≥8 is considered high-quality, 5-7 is considered moderate-quality, and ≤4 is considered low-quality.

2.3.1.6 Reviews using Fors' risk of bias scale

Fors et al., 2011

In their systematic review, Fors et al. (2011) evaluated 18 RCTs of psychological treatments across numerous outcomes in BCa. Seventeen of the 18 RCTs included

emotional distress as an outcome. Fourteen of those were categorised as moderate-quality and three were categorised as high-quality. In most RCTs, adequate methods of randomisation were employed, reliable outcome measures were used, and statistical methods appropriately tested the hypotheses. However, most RCTs failed to blind participants, personnel or outcome assessors.

2.3.1.7 Naaman's quality scale

Naaman et al. (2009) developed their own quality scale based on a framework provided by Cook & Campbell (1979) and the Jadad scale (Jadad et al., 1996). It consists of seven items, each scored 0 (no) or 1 (yes), producing a total score ranging from 0-7. Three of the items assess two bias domains (random sequence generation and incomplete outcome data) and four assess other aspects of quality (power analysis, control for patient demoralisation, manualisation, and therapist's adherence to treatment). An RCT with a score of ≥5 is considered high-quality and <5 is considered low-quality.

2.3.1.8 Reviews using Naaman's quality scale

Naaman et al., 2009

Naaman and colleagues (2009) meta-analysis evaluated 18 RCTs of psychological treatments for emotional distress and QoL in BCa. Sixteen of the 18 RCTs included emotional distress as an outcome. Seven of those RCTs were rated as high-quality and nine as low-quality. Naaman et al. (2009) only reported whether an RCT was of 'high' or 'low' quality. The authors did not report the overall quality score of each RCT or provide scores for individual items, making it impossible to know which methodological aspects were well-conducted or most in need of improvement.

2.3.1.9 Beatty's quality checklist

Beatty et al. (2018) created their own quality checklist which combined the five criteria for empirically supported psychotherapies (type of comparator condition, sample size, power analysis, reliable outcome measures, and clear specification of inclusion criteria; Chambless & Hollon, 1998) and six domains from the Cochrane RoB tool (random sequence generation, allocation concealment, blinding of participants and personnel, incomplete outcome data, selective outcome reporting, and

'other biases'; Higgins & Green, 2011). Beatty et al. (2018) defined 'higher-quality trials' as those meeting at least 8 of the 11 criteria.

2.3.1.10 Review using Beatty's quality checklist

Beatty et al., 2018

In their systematic review, Beatty et al. (2018) evaluated 15 RCTs of psychological treatments for emotional distress, survival, QoL, and somatic symptoms in metastatic BCa patients. Thirteen of the 15 RCTs included emotional distress as an outcome. Eight of those were categorised as 'higher-quality' and five as 'lower-quality'. Most RCTs had an adequate sample size, clearly specified inclusion criteria, reliable outcome measures, adequate methods of sequence generation and allocation concealment, were free of selective reporting, and used 'acceptable' comparator conditions (though 'acceptable' was not defined). Most RCTs failed to conduct power analysis, blind participants and personnel, or remain free from 'other biases'; and just under half reported strategies to deal with incomplete outcome data.

2.3.2 Summary of results

Seven of the nine reviews assessed RoB and two assessed the broader domain of quality. However, the seven reviews assessing RoB claimed to assess the broader domain of quality. The design elements assessed across the different assessment tools varied greatly. For example, only 27% of the design elements included in Beatty's quality checklist (Beatty et al., 2018) were included in Naaman's quality scale (Naaman et al., 2009). The methodological quality of RCTs varied drastically across reviews. This was in part due to the different RCTs included in reviews but mainly due to the different design elements being assessed. For example, Mustafa et al. (2013) and Beatty et al (2013) assessed the same eight RCTs using two different assessment tools. Mustafa et al. (2013) categorised all eight as having an unclear RoB, while Beatty et al. (2018) categorised five as 'higher quality', and three as 'lower quality'.

Four reviews used assessment tools with well-defined criteria (Haller et al., 2017; Jassim et al., 2015; Mustafa et al., 2013; Ye et al., 2018) but five (including both reviews assessing the broader construct of quality) used assessment tools with poorly operationally defined criteria (Beatty et al., 2018; Fors et al., 2011; Naaman et al., 2009; Xiao et al., 2017; Zhang et al., 2016). For example, to assess blinding of

participants and personnel, Fors RoB scale asks: *'are blinding methods adequate?'* without providing an explanation of what constitutes '*adequate'*. Seven reviews reported ratings for each individual design element (Beatty et al., 2018; Fors et al., 2011; Haller et al., 2017; Jassim et al., 2015; Mustafa et al., 2013; Ye et al., 2018; Zhang et al., 2016) but two stated only the overall quality of RCTs (Naaman et al., 2009; Xiao et al., 2017).

Table 2.2: Trial quality findings from each review

Review	Risk of bias or quality assessment	Type of assessment tool	Assessment tool	RCTs	Quality of RCTs
Beatty et al. 2018	Quality	Checklist	Beatty's Quality checklist	n=13	Higher quality n=8 (62%) Lower quality n=5 (38%)
Fors et al. 2017	Risk of bias	Scale	Fors' RoB scale	n=17	High quality n=3 (18%) Moderate quality n=14 (82%)
Haller et al. 2017	Risk of bias	Domain based evaluation	Cochrane RoB tool	n=9	High RoB n=9 (100%)
Jassim et al. 2015	Risk of bias	Domain based evaluation	Cochrane RoB tool	n=28	High RoB n=9 (32%) Unclear RoB n=19 (68%)
Mustafa et al. 2015	Risk of bias	Domain based evaluation	Cochrane RoB tool	n=8	Unclear RoB n=8 (100%)
Naaman et al. 2009	Quality	Scale	Naaman's Quality scale	n=16	High quality n=7 (44%) Low quality n=9 (56%)
Xiao et al. 2017	Risk of bias	Scale	Jadad scale	n=13	Adequate quality n=13 (100%)
Ye et al. 2018	Risk of bias	Domain based evaluation	Cochrane RoB tool	n=10	High RoB n=8 (80%) Unclear RoB n=1 (10%) Low RoB n=1 (10%)
Zhang et al. 2016	Risk of bias	Scale	Jadad scale	n=5	High quality n=2 (40%) Adequate quality n=1 (20%) Low quality n=2 (40%)

Note. RoB = risk of bias; RCTs = randomised controlled trials

2.4 Discussion

A scoping review of reviews assessing the methodological quality of RCTs for emotional distress in BCa was conducted. Nine reviews were identified, in which five different assessment tools were used. The methodological quality of RCTs included in these reviews varied drastically depending on the design elements being assessed.

Seven of the nine reviews took a narrow approach to assessing methodological quality by only focusing on internal validity (i.e. RoB). However, all seven claimed to assess the broader domain of methodological quality. The remaining two reviews (corresponding to two assessment tools) went beyond assessing RoB and assessed other important design elements essential to high-quality RCTs (i.e. use of power analysis, type of comparator conditions, and use of reliable outcome measures). Yet, they still failed to assess many other design elements equally essential to high-quality RCTs (i.e. clarity of sample description, representativeness of the sample, validity and specificity of outcome measures, control of concomitant treatments, length of follow-up, and assessment of clinical significance).

Only one of the previous reviews (corresponding to one assessment tool) assessed any design elements specific to psychotherapy RCTs (Naaman et al. 2009). Naaman et al. (2009) assessed manualisation and therapist's adherence. However, none of the reviews (or the assessment tools used) assessed the number of therapists, therapist's level of training, therapist's competence to treatment, or equality of therapy hours between conditions. Moreover, Naaman et al. (2009) only reported whether a trial was of 'high' or 'low' quality. Thus, the extent to which manuals were followed or therapists adhered to treatment remains unknown.

Finally, the two reviews that assessed the broader construct of quality used assessment tools that were poorly operationally defined. This increases the likelihood of heterogenous ratings across different raters, limiting the reliability and validity of any conclusions drawn.

2.4.1 Conclusion

The methodological quality of RCTs of psychological treatments for emotional distress in BCa has been inadequately assessed. Unless the quality of these RCTs is adequately assessed, confidence in conclusions regarding treatment efficacy will be

undermined. An assessment of the methodological quality of RCTs of psychological treatments for emotional distress in BCa using an operationally defined assessment tool which assesses important generic design elements and psychotherapy-specific design elements is needed.

Therefore, in the next chapter, a systematic review of the methodological quality of RCTs of psychological treatments for emotional distress in BCa is presented.

Chapter 3. Study 2: A Systematic Review of the Quality of Randomised Controlled Trials of Psychological Treatments for Emotional Distress in Breast Cancer

3.1 Introduction

The aim of this study was to evaluate the methodological quality of RCTs of psychological treatments for emotional distress in BCa using a scale developed in the context of broader psychotherapy research. The most appropriate such scale is the Psychotherapy Outcome Study Methodology Rating Form (POMRF; Öst, 2008), which was explicitly designed to assess the quality of psychotherapy RCTs. It assesses both generic design elements and those specific to psychotherapy RCTs. The POMRF has been used to assess the quality of psychotherapy RCTs in several mental health populations (Arnberg & Öst, 2014; Öst, 2008, 2014; Öst et al., 2015; Öst & Ollendick, 2017; Öst, Riise, Wergeland, Hansen, & Kvale, 2016; Sloan et al., 2017; Swain, Hancock, Dixon, & Bowman, 2015; Swain, Hancock, Hainsworth, & Bowman, 2013), and those assessments provide a benchmark against which to gauge the quality of RCTs in BCa. Therefore, using the POMRF, this study had five aims:

Aim 1: Evaluate the overall quality of RCTs of psychological treatments for emotional distress in BCa, considering both generic design elements and those specific to psychotherapy RCTs.

Aim 2: Evaluate specific design elements that have previously been inadequately evaluated in meta-analyses or are poorly implemented in RCTs.

Aim 3: Assess the quality of RCTs in this population against the benchmark of RCTs in mental health populations.

Aim 4: Assess whether the quality of RCTs differ depending on the type of treatment being tested.

Aim 5: Considering the general improvement of methodological standards in psychotherapy RCTs over time, determine whether methodological quality in BCa studies has improved over time.

3.2 Method

3.2.1 Search strategy

PubMed, PsycINFO, Web of Science, Scopus, PsycARTICLE, and AMED were searched from their inception until October 2016 using MeSH terms and keywords to

identify psychotherapy RCTs for emotional distress in BCa. The search used a combination of terms associated with BCa (including: "*breast neoplasms" OR "breast cancer"),* emotional distress (including: "*Depression"[Mesh] OR "depressive disorder"[Mesh] OR "depressive disorder" OR "anxiety"[Mesh] Or "anxiety disorders"[Mesh] OR "anxiety disorders" OR "anxiety" OR "depression" OR "emotional distress" OR "psychological distress"*), and psychological treatments (including: *"Psychotherapy"[Mesh] OR "psychotherapy" OR "psychological therapy" OR "counselling" OR "counselling" OR "psychological intervention" OR "cognitive behavioural therapy" OR "group therapy" OR "psychosocial therapy"*). No date restriction was applied but only English language articles were included. The final search strategy used for each database is available in Appendix 2. In addition, relevant meta-analyses and reference lists of eligible articles were hand-searched to identify any additional studies that may have been missed.

3.2.2 Eligibility criteria

Eligibility criteria are detailed according to the PICOS framework (Liberati et al., 2009).

Participants: Trials in which the included participants were exclusively adults aged 18 years or older with a histologically confirmed diagnosis of BCa.

Interventions: Trials evaluating a psychological treatment. As "psychological treatment" is poorly defined in the literature (Jassim et al., 2015) a generic definition was used: treatments primarily using psychological techniques. Trials evaluating complementary alternative medicines (i.e. yoga, hypnosis, reiki, logotherapy, art therapy, dance therapy) or treatments involving no interaction between therapist and patient (i.e. based explicitly on written or visual material) were excluded.

Controls: Trials using either a no treatment (usual care) control, an active (attention placebo) control or an alternative psychological treatment.

Outcomes: Trials in which the primary and/or secondary outcome was emotional distress, defined as anxiety, depression, general mood, or global emotional distress.

Studies: Only RCTs published in English in a peer-reviewed journal.

3.2.3 Study selection

After removing duplicates, titles and abstracts were searched to identify relevant RCTs by one reviewer (JT). If this information could not be ascertained from the title or abstract, the full text of the article was obtained for detailed scrutiny. Any uncertainties regarding trial inclusion were discussed with a second reviewer (PF). When a single trial was published more than once, the report that most thoroughly presented the methods and findings was evaluated. Therefore, each paper represented a unique trial.

3.2.4 Data extraction

Using a standardised data extraction protocol (see Appendix 3), data were extracted for all included trials by two independent reviewers (JT & CH). Data extracted was year of publication; country of origin; number of participants randomly assigned to condition; mean age; distribution of trajectory stage; distribution of tumour stage; outcome measures; treatment type; treatment format; duration of treatment; number of treatment sessions; and type of control condition.

3.2.5 Aim 1: Overall quality of trials

Methodological quality was rated using the POMRF (see Appendix 4; Öst, 2008). It consists of 22 items, each scored 0 (poor), 1 (fair), or 2 (good), producing a total score ranging from 0 to 44, with higher scores indicating greater quality. Three items relating to psychiatric diagnoses (items 2, 4, & 8), irrelevant to this review, were disregarded; therefore, in this study, the maximum possible score was 38. A minimum cut-off score to determine adequate methodological quality on the POMRF has not been established. However, a review (Gerber et al., 2011) which used the Randomized Controlled Trial Psychotherapy Quality Rating Scale (RCT-PQRS; Kocsis et al., 2010) to evaluate the quality of psychodynamic trials provided a suitable benchmark. In that review, a cut off score of at least 50% of the maximum possible score on the RCT-PQRS was used. Thus, in this review a total score of 19 out of 38 (i.e. 50% of the maximum possible score) was chosen as the criterion for minimum adequate quality.

To compare quality on generic and psychotherapy-specific items, POMRF items were allocated to two subscales: “generic design elements” (Table 3.1: maximum possible score of 26) and “psychotherapy-specific design elements” (Table 3.1: maximum

possible score of 12). To allow comparison between the two subscales, total subscale scores were transformed into percentages of the maximum possible on the relevant subscale.

Two reviewers (JT & CH) independently rated the quality of each trial. To determine consistency of quality scores between the reviewers, interrater reliability was assessed using the intra-class correlation coefficient (ICC) for total quality scores, and the weighted kappa statistic for individual item scores. The ICC for total quality scores was 0.95 (95% CI, 0.92–0.97) and kappa for individual items ranged from 0.73 to 0.93, with mean 0.8, indicating good inter-rater reliability. Following the assessment of inter-rater reliability, discrepancies in ratings were resolved through discussion and consensus between both reviewers.

3.2.6 Aim 2: Quality of specific design elements

All design elements specific to psychotherapy RCTs (Table 3.1) and generic ones that were particularly poorly implemented (i.e. a score of zero in at least 75% of trials) were descriptively evaluated.

3.2.7 Aim 3: Quality comparison with mental health populations

To locate meta-analyses and systematic reviews evaluating the quality of RCTs using the POMRF in mental health populations, all papers citing the study in which the POMRF was devised were identified by searching Google Scholar. Potentially relevant papers were retrieved and assessed for eligibility. To compare the quality of RCTs in BCa with RCTs in mental health populations, for which the full 22-item scale was reported, scores were transformed into percentages of the maximum possible score on the scale, as well as on the two subscales.

Table 3.1: Items in the Psychotherapy outcome study methodology rating form

Subscale 1: Generic design elements (maximum possible score of 26)	Subscale 2: Psychotherapy-specific design elements (maximum possible score of 12)
1. Clarity of sample description	13. Manualised treatment
3. Representativeness of sample	14. Number of therapists
5. Specificity of outcome measures	15. Therapist training/experience
6. Reliability and validity of outcome measures	16. Checks for therapist adherence
7. Use of blind evaluators	17. Checks for therapist competence
9. Assignment to treatment	22. Equality of therapy hours
10. Design (i.e. type of comparator condition)	
11. Power analysis	
12. Assessment points	
18. Control of concomitant treatments	
19. Handling of attrition	
20. Statistical analyses and presentation of results	
21. Clinical significance	

3.2.8 Aim 4: Quality comparison by treatment type

Psychological treatments were coded into five categories: *'cognitive-behavioural-based treatments'* (CBT; treatments targeting specific thoughts or behaviours using cognitive behavioural techniques); *'mindfulness-based therapies'* (MBT; treatments focusing on meditation, visualisation, and present-moment awareness); *'Psycho-education'* (treatments primarily providing psychological education) *'support'* (treatments emphasising a supportive environment by providing emotional or social support); and '*peer-led treatments'* (any treatment that was delivered by non-professional peers). If a treatment did not fit into one of the five categories (i.e. it combined components from multiple categories without emphasising any one) it was

categorised into a sixth category - *'other'.* The categorisation of treatments was discussed by three reviewers (JT, CH & PF) until consensus was reached.

One-way ANOVAs compared differences in trial quality between treatment categories. Total POMRF scores and the two subscales scores distinguishing generic and psychotherapy-specific design elements were evaluated. Post hoc Tukey HSD tests followed significant main effects to identify which treatment types differed.

3.2.9 Aim 5: Quality trends over time

Spearman correlation was calculated for year of study publication with total POMRF scores and the two subscale scores distinguishing generic and psychotherapy-specific design elements.

3.3 Results

The electronic database search yielded 2,210 citations; an additional 18 were identified through hand searching. After removal of duplicates, 1,412 remained for screening based on title and abstract. Of these, 1,169 clearly did not meet the inclusion criteria. The full text articles of the remaining 252 citations were retrieved and assessed. Ninety-one articles published from 1980 through October 2016 were eligible and included. Figure 3.1 shows the study selection process. A complete list of references of the included RCTs can be found in Appendix 5.

A summary of the trial characteristics is presented in Table 3.2, while a complete description of each trial is presented in Table 3.3. The trials comprised 13,553 patients with sample sizes ranging from 14-558 (mean 149; median 117). Most trials were conducted in the United States and exclusively included non-metastatic BCa patients. The treatment approach used most frequently was CBT and most treatments were delivered in group format. On average, treatment involved a mean of 8 sessions (median 8; range 1-25), with each session lasting 1.5 hours (median 1.5; range 0.25-3; excluding supplementary material and subsequent monitoring). Depression was the most common outcome variable (n=60), followed by anxiety (n=47) and mood/global distress (n=44). The Profile of Mood States (POMS; n=20), Centre for Epidemiological Studies Depression Scale (CES-D; n=19) and the Hospital Anxiety and Depression scale (HADS; n=17) were the most commonly used outcome measures.

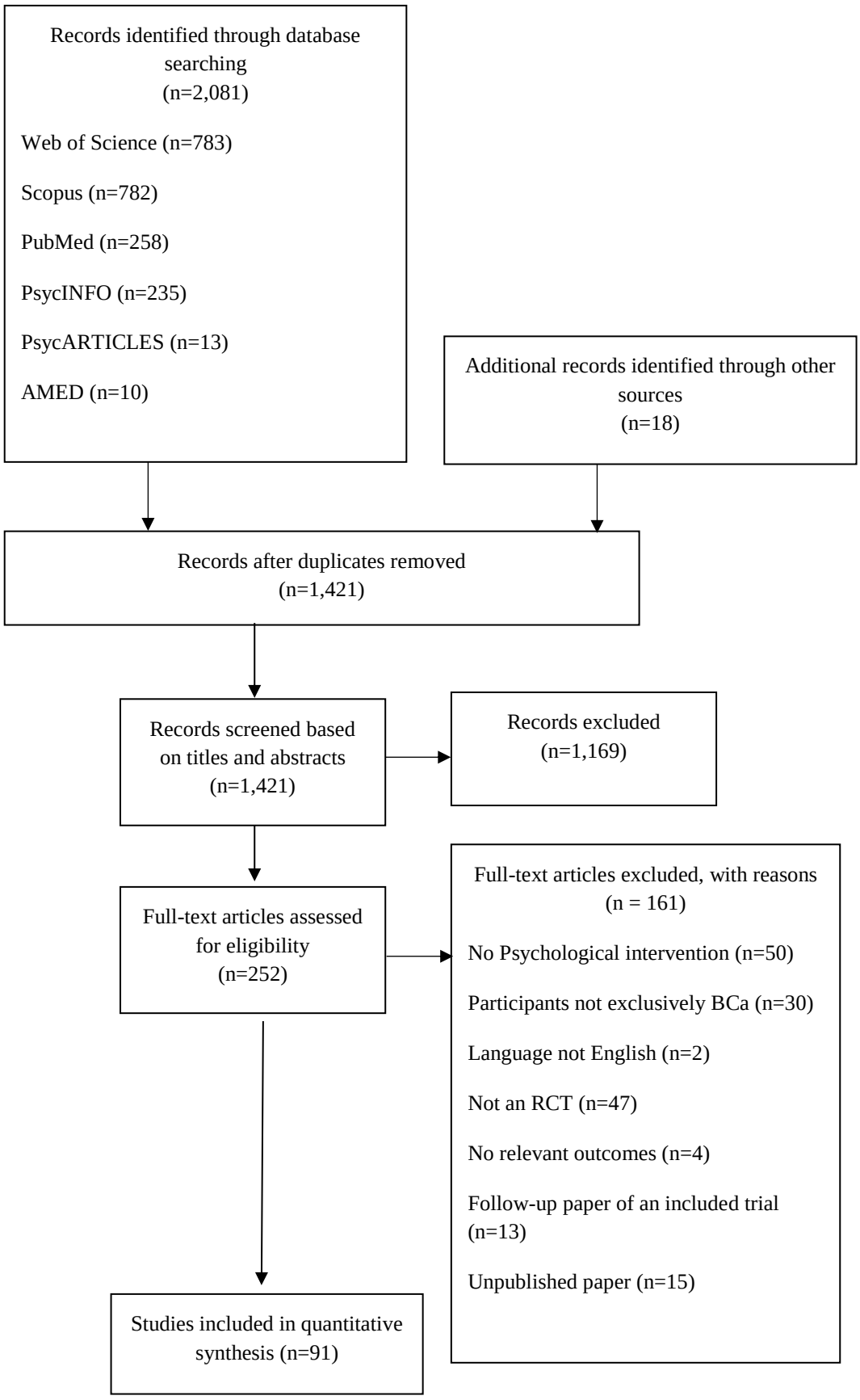


Figure 3.1: Flow chart showing trial identification and selection

Table 3.2: Descriptive summary of included trials by treatment type and format

			Treatment format						Treatment type											
	Total sample (n=91)		**Individual (n=37)**		**Group (n=48)**		**Couples (n=6)**		**CBT (n=40)**		**Mindfulness (n=5)**		**PsyEd (n=6)**		**Support (n=21)**		**Peer-led (n=6)**		**Other (n=13)**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Population																				
Patients																				
Total No. of patients	13,553		5,553		7,334		666		4,809		922		1,102		3,913		1,234		1,573	
Mean sample size per study	149		150		153		111		120		184		184		186		205		121	
Median sample size per study	117		120		119		46		100		172		162		152		198		87	
Minimum sample size	14		25		32		14		14		71		66		46		104		40	
Maximum sample size	558		558		382		302		355		366		367		558		305		382	
Mean age, years	52		53		52		50		52		52		49		54		52		52	
Median age, years	52		54		51		52		53		50		50		53		51		53	
Stage of disease																				
Non-metastatic	59	65%	23	62%	31	65%	5	83%	28	70%	5	100%	5	83%	11	52%	3	50%	7	54%

			Treatment format						Treatment type											
	Total sample (n=91)		**Individual (n=37)**		**Group (n=48)**		**Couples (n=6)**		**CBT (n=40)**		**Mindfulness (n=5)**		**PsyEd (n=6)**		**Support (n=21)**		**Peer-led (n=6)**		**Other (n=13)**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Metastatic	7	8%	1	3%	6	13%			3	8%					4	19%			3	23%
Both	15	16%	9	24%	5	10%	1	17%	6	15%					4	19%	2	33%		
NR	10	11%	4	11%	6	13%			3	8%			1	17%	2	10%	1	17%	3	23%
Country																				
USA	45	49%	20	54%	20	42%	5	83%	21	53%	3	60%	3	50%	10	48%	5	83%	3	23%
Canada	10	11%	4	11%	6	13%			6	15%			1	17%	2	10%			1	8%
Australia	7	8%	2	5%	5	10%			3	8%			1	17%	2	10%			1	8%
UK	4	4%	3	8%	1	2%					1	20%			3	14%				
Sweden	4	4%	2	5%	2	4%			2	5%					1	5%			1	8%
Other	21	23%	5	14%	13	27%	1	17%	8	20%	1	20%	1	17%	3	14%	1	17%	7	54%
	China (2), Croatia (1), Denmark (2), France (2), Germany (1), Greece (2), Holland (1), Iran (2), Ireland (1), Israel (1), Italy (1), Japan (2), Korea (1), Norway (1), Romania (1)		Croatia (1), Germany (1), Italy (1), Korea (1), Romania (1), Greece (1)		China (2), Denmark (2), France (2), Holland (1), Iran (1), Ireland (2), Israel (1), Japan (2), Norway (1)		Greece (1)		China (1), Croatia (1), Denmark (1), France (1), Iran (1), Ireland (1), Israel (1), Italy (1)		Denmark (1)		Norway (1)		China (1), Japan (1), Romania (1)		Korea (1)		France (1), Germany (1), Greece (2), Holland (1), Iran (1), Japan (1),	
Exclusively distressed patients																				
Yes	12	13%	7	19%	5	10%			8	20%					1	5%	1	17%	2	15%

			Treatment format						Treatment type											
	Total sample (n=91)		**Individual (n=37)**		**Group (n=48)**		**Couples (n=6)**		**CBT (n=40)**		**Mindfulness (n=5)**		**PsyEd (n=6)**		**Support (n=21)**		**Peer-led (n=6)**		**Other (n=13)**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
No	79	87%	29	81%	43	90%	6	100%	32	80%	5	100%	6	100%	20	95%	5	83%	11	85%
Outcomes *																				
Outcome variable																				
Anxiety	47	52%	22	60%	22	46%	3	50%	24	60%	3	60%	3	50%	9	43%	2	33%	6	46%
Depression	60	66%	29	78%	27	56%	4	67%	26	65%	4	80%	4	67%	12	57%	5	83%	9	69%
Mood/ global distress	44	48%	10	27%	30	63%	4	67%	21	53%	2	40%	3	50%	10	48%	2	33%	6	46%
Outcome measure																				
CES-D	20	22%																		
POMS	19	21%																		
HADS	17	18%																		
Treatment (active treatment)																				
No. of sessions																				
Mean	8		7		9		6		9		8		5		7		7		9	
Median	8		6		9		6													
Minimum	1		1		1		4		1		6		3		1		3		2	

	Total sample (n=91)		Treatment format: Individual (n=37)		Group (n=48)		Couples (n=6)		Treatment type: CBT (n=40)		Mindfulness (n=5)		PsyEd (n=6)		Support (n=21)		Peer-led (n=6)		Other (n=13)	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Maximum	25		25		23		8		23		11		6		16		9		25	
Variable No. of sessions (trials)	9		4		5				2						5		2			
Not reported (trials)	3		3						2						1					
Length of sessions (hours)																				
Mean	1.5		0.75		1.75		0.75		1.5		2		1.5		1.25		1.25		1.5	
Median	1.5		0.75		2		1.5				2									
Minimum	0.25		0.25		1		0.5		0.5		2		0.25		0.5		1		0.5	
Maximum	3		1.5		3		1.5		3		2		2.5		2		1.5		2.5	
Variable no of sessions (trials)	12		6		6				4		3				3		2			
Not reported (trials)	14		8		4		2								3		1		4	
Treatment type																				
CBT	40	44%	14	38%	22	46%	4	67%												
Mindfulness	5	5%			5	10%														

	Treatment format								Treatment type											
	Total sample (n=91)		**Individual (n=37)**		**Group (n=48)**		**Couples (n=6)**		**CBT (n=40)**		**Mindfulness (n=5)**		**PsyEd (n=6)**		**Support (n=21)**		**Peer-led (n=6)**		**Other (n=13)**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
PsyEd	6	7%	2	5%	4	8%														
Support	21	23%	11	30%	9	19%	1	17%												
Peer-led	6	7%	5	14%	1	2%														
Other	13	14%	5	14%	7	15%	1	17%												
Treatment Format																				
Individual	37	41%							14	35%			2	33%	11	52%	5	83%	5	39%
Group	48	53%							22	55%	5	100%	4	67%	9	43%	1	17%	7	54%
Couples	6	7%							4	10%					1	5%			1	8%

*Note***.** No. = number; NR = not reported; CBT = cognitive behavioural therapy; PsyEd = psychoeducation; * Because several trials used multiple outcome measures, the number of trials presented for the type of outcome measure exceeds the total number of trials

Table 3.3: Descriptive summary of each included trial

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures
Abad 2016	Iran	36	NR	NR	Cognitive behavioural	CBT	Ind	8	2 hrs	No	Anx, dep	DASS
Allard 2007	Canada	117	54	I-IIa	Attention focused, symptom management, coping mechanisms	Other	Ind	2	NR	No	Mood/ distress	POMS-SF
Allen 2002	USA	164	42	I-IIIa	Problem solving training	CBT	Ind	6	NR	No	Mood/ distress	MHI-5, IES
Andersen 2004	USA	227	51	II-III	Coping skills, stress management, relaxation, health behaviour, communication	Other	Grp	18	1.5 hrs	No	Mood/ distress	POMS
Anton 2015	Croatia	120	57	NR	Cognitive behavioural	CBT	Ind	Varied	Varied	Yes	Anx, dep	HAM-D, HAM-A
Antoni 2001	USA	136	50	0-II	CBSM	CBT	Grp	10	2 hrs	No	Dep, mood/ distress	CES-D, POMS composite
Antoni 2006	USA	199	50	0-III	CBSM	CBT	Grp	10	2 hrs	No	Anx, mood/ distress	ABS composite, HAM-A
Antoni 2009	USA	128	50	0-III	CBSM	CBT	Grp	10	NR	No	Anx, dep, mood/ distress	ABS composite, HAM-A
Arving 2007	Sweden	179	55	0-IV	Relaxation, distraction, techniques, problem solving	CBT	Ind	0-23	35 mins - 1hr	No	Anx, dep	HADS, STAI

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures
Ashing 2014	USA	211	53	0-III	Education; problem solving training; stress management;	CBT	Ind	8	40-50 mins	Yes	Dep	CES-D
Badger 2007	USA	96	54	I-III	Interpersonal counselling	Support	Ind	6	35 mins	No	Anx, dep	CES-D, self-constructed anxiety measure
Badger 2013	USA	52	52	I-IV	Interpersonal counselling	Support	Cpl	8	30 mins	No	Anx, dep	CES-D, STAI, PANAS-Negative affect subscale
Badger 2013	USA	90	47	I-III	Interpersonal counselling	Support	Ind	8	29 mins	No	Dep	CES-D
Baucom 2009	USA	14	NR	I-II	Relationship enhancement, problem solving skills, emotional expression	CBT	Cpl	6	1.25 hrs	No	Mood/ distress	BSI-18
Beutel 2014	Germany	157	52	0-IV	Psychodynamic	Other	Ind	25	NR	Yes	Dep	HADS-D, SCID
Bjornklett 2012	Sweden	382	58	NR	Education, support, relaxation, Qi-gong, liberating dance	Other	Grp	7 + 4 follow-up	NR	No	Anx, dep	HADS

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures
Boesen 2011	Denmark	205	NR	I-IIIa	Existential–cognitive therapy	CBT	Grp	10	32 hrs	No	Mood/ distress	POMS-SF
Bower 2015	USA	71	47	0-III	Mindfulness meditation	Mindfulness	Grp	6	2 hrs	No	Dep	CES-D
Bredal 2014	Norway	367	55	0-III	Education, stress management, support	PsyEd	Grp	5	2 hrs	No	Anx, dep	HADS
Budin 2008	USA	249	54	0-III	Support, counselling, education,	Support	Ind	8	NR	No	Mood/ distress	PAL-C psychological sub-scale
Burton 1995	UK	215	62	I-IV	Counselling	Support	Ind	1	NR	No	Anx, dep	HADS
Carlson 2013	Canada	271	55	I-IV	SET	Support	Grp	8 (+ 1 workshop)	1.5 hrs (+ 6 hrs)	Yes	Mood/ distress	POMS
Chan 2006	China	87	49	I-III	SET	Support	Grp	8	2 hrs	No	Mood/ distress	GHQ-12
Christensen 1983	USA	20	40	0-III	Problem solving, communication, emotional expression	Other	Cpl	4	NR	No	Anx, dep	BDI, STAI-state
Classen 2001	USA	125	53	IV	SET	Support	Grp	Varied	1.5 hrs	No	Mood/distress	POMS
Classen 2008	USA	353	50	I-IIIa	SET	Support	Grp	12	1.5 hrs	No	Anx, dep, mood/ distress	POMS, HADS

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures
Cohen 2007	Israel	114	54	I-II	Cognitive behavioural	CBT	Grp	9	1.5 hrs	No	Mood/ distress	BSI
Cousson-Gelie 2011	France	92	53	0-III	Managing irrational thoughts, perceived control, coping, relaxation	Other	Grp	8	2 hrs	No	Anx, dep	HADS
Creuss 2000	USA	34	46	I-II	CBSM	CBT	Grp	10	2 hrs	No	Mood/distress	POMS
Davis 1986	Canada	25	51	I	Cognitive	CBT	Ind	13	45 mins	No	Anx	STAI-state
Dirksen 2007	USA	81	58	I-III	Cognitive behavioural	CBT	Grp	6	Varied	No	Anx, dep	CES-D, STAI
Dolbeault 2009	France	203	53	NR	Problem solving, cognitive restructuring, education	CBT	Grp	8	2 hrs	No	Anx, mood/ distress	POMS, STAI
Dowlatabadi 2016	Iran	42	37	NR	Positive psychotherapy	Other	Grp	10	1.5 hrs	Yes	Dep	BDI
Edelman 1999[a]	Australia	60	48	I-II	Cognitive behavioural	CBT	Grp	12	2 hrs	No	Anx, dep	POMS
Edelman 1999[b]	Australia	124	50	IV	Cognitive behavioural	CBT	Grp	11	NR	No	Mood/ distress	POMS
Edmonds 1999	Canada	66	51	IV	CBT, coping skills training	CBT	Grp	35 (+15 additional if requested)	2 hrs	No	Mood/ distress	POMS, POMS-SF

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures
Ferguson 2012	USA	40	50	I-IIIa	Cognitive behavioural	CBT	Ind	4	30-50 mins	No	Anx, dep	CES-D, STAI
Ferguson 2016	USA	47	55	I-III	Cognitive behavioural	CBT	Ind	8	30-45 mins	Yes	Anx, dep	DASS-21
Fillion 2008	Canada	94	52	0-III	Education, stress management, exercise	PsyEd	Grp	4 (+1 booster)	2.5 hrs (+ 15 mins)	No	Mood/ distress	POMS
Fukui 2000	Japan	50	53	I-III	Coping skills training, CBT, education	Other	Grp	6	1.5 hrs	No	Anx, mood/ distress	HADS, POMS
Gaston 2000	USA	110	NR	II-IV	Cognitive restructuring, relaxation/guided imagery	CBT	Ind	1	NR	No	Anx, dep	BDI, STAI
Giese-Davis 2016	USA	104	NR	0-IV	Peer counselling, support	Support	Ind	Varied	Varied	No	Dep	CES-D
Goodwin 2001	Canada	235	51	IV	SET	Support	Grp	>52	1.5 hrs	No	Mood/ distress	POMS
Gotay 2007	USA	305	54	I-IIIa	Peer counselling, support, stress management	Support	Ind	04-Aug	NR	No	Dep	CES-D
Graves 2003	USA	32	56	0-III	Coping skills training, relaxation, CBT	CBT	Grp	8	1.5 hrs	No	Mood/ distress	POMS
Groarke 2012	Ireland	355	54	0-IV	CBSM	CBT	Grp	5	3 hrs	No	Anx, dep	HADS

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures
Gudenkauf 2015	USA	183	54	0-III	CBSM	CBT	Grp	5	1.5 hrs	No	Dep	ABS depression subscale, IES-R
Heiney 2003	USA	66	50	NR	Education, coping strategies, support	PsyEd	Grp	6	1.5 hrs	No	Mood/ distress	POMS-SF
Henderson 2012	USA	172	50	I-II	MBSR	Mindfulness	Grp	7 + (1 retreat session + 3 booster)	3.5 hrs (+7.5 +2 hrs)	No	Anx, mood/ distress	BDI, BAI, SCL-90-R
Hoffman 2012	UK	229	50	0-III	MBSR	Mindfulness	Grp	8 (+ 1 extra)	2 hrs (+ 6 hrs)	No	Mood/ distress	POMS
Hopko 2011	USA	80	55	0-IV	BATD	CBT	Ind	8	1 hr	Yes	Anx, dep	BDI, HAM-D, BAI
Kalaitzi 2007	Greece	40	53	0	Couple and sex therapy	Other	Cpl	6	NR	No	Mood/ distress	CES-D, STAI
Kissane 2003	Australia	303	46	I-II	Cognitive existential	CBT	Grp	20 (+3 relaxation)	1.5 hrs (+50 mins)	No	Anx, mood/ distress	ABS, HADS
Kissane 2007	Australia	227	52	IV	SET	Support	Grp	Varied	1.5 hrs	No	Mood/ distress	IES
Lane 2005	Australia	42	54	NR	Personal construct therapy	Other	Grp	9	2 hrs	No	Anx, dep	Gottschalk-Glaser content

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures
												analysis scale
Larson 2000	USA	41	56	I-IV	Education, problem-solving, relaxation, support	Other	Ind	2	1.5 hrs	No	Dep	CES-D
Lechner 2014	USA	114	51	0-IV	CBSM	CBT	Grp	10	1.5 hrs	Yes	Dep, mood/ distress	CES-D, POMS-SF
Lee 2013	Korea	129	48	I-III	Peer support	Support	Ind	6	Varied	Yes	Anx, dep	HADS
Lengacher 2009	USA	84	58	0-III	MBSR	Minfdulness	Grp	6	2 hrs	No	Anx, dep	CES-D, STAI
Lewis 2015	USA	213	43	0-III	Education, counselling	PsyEd	Ind	5	1.5 hrs	No	Anx, dep	CES-D
Maguire 1980	UK	152	NR	NR	Counselling	Support	Ind	NR	NR	No	Anx, dep	Present state examination
Manne 2005	USA	238	50	0-IIIa	Communication, relaxation, stress management, problem solving	CBT	Cpl	6	1.5 hrs	No	Anx, mood/ distress	MHI-18. IES
Manne 2016	USA	302	55	0-IIIa	Communication, relaxation, stress management, problem solving, cognitive restructuring	CBT	Cpl	8	1.5 hrs	No	Anx, mood/ distress	MHI-38

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures
Marchioro 1996	Italy	36	NR	0-III	Cognitive behavioural	CBT	Ind	NR	50 mins	No	Dep	BDI
Marcus 2010	USA	304	NR	I-IIIa	Counselling, relaxation	Support	Ind	16	45 mins	No	Dep	CES-D
McArdle 1996	UK	272	NR	NR	Support, counselling	Support	Ind	Varied	Varied	No	Anx, dep	HADS
Mens 2015	USA	245	51	I-II or IV	Peer support	Support	Grp	8	1 hrs	No	Dep	CES-D
Miyashita 2005	Japan	78	51	0-III	SET	Support	Grp	1	1.5-2 hrs	No	Anx	STAI-state
Napoles 2015	USA	151	51	0-III	CBSM	CBT	Ind	9	1.5 hrs	No	Anx, mood/ distress	BSI
Naumann 2012	Australia	46	51	I-III	Support, counselling, exercise	Support	Ind	8	1 hrs	No	Dep	BDI
Pelekasis 2016	Greece	61	56	I-IV	CBT, guided imagery, dietary consulting, exercise,	Other	Ind	6	30 mins	No	Anx, dep	DASS-21
Qiu 2013	China	62	51	0-IV	Cognitive behavioural	CBT	Grp	10	2 hrs	Yes	Anx, dep	HAM-D, SAS
Rissanen 2015	Sweden	155	58	I-III	CBSM	CBT	Grp	10	2 hrs	Yes	Anx, dep, mood/ distress	HADS, IES

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures
Samarel 1997	USA	181	53	I-II	Support, stress management, communication, problem solving	Support	Grp	8	2 hrs	No	Mood/ distress	POMS-LASA
Sandgren 2000	USA	62	52	I-II	Support, problem solving, cognitive restructuring	CBT	Ind	10	30 mins	No	Mood/ distress	POMS
Sandgren 2003	USA	235	55	I-III	Education	PsyEd	Ind	5 (+1 booster)	30 mins	No	Anx, dep, mood/ distress	POMS
Savard 2005	Canada	54	54	I-III	Cognitive behavioural	CBT	Grp	8 (+1 optional booster)	1.5 hrs	No	Anx, dep	HADS
Savard 2006	Canada	45	52	IV	Cognitive behavioural	CBT	Ind	8 (+ 3 booster)	1-1.5 hrs	Yes	Anx, dep	BDI, HAM-D, HADS
Savard 2014	Canada	242	54	0-III	Cognitive behavioural	CBT	Ind	6	50 mins	No	Anx, dep	HADS
Scheier 2005	USA	252	44	0-II	Education	PsyEd	Grp	4	2 hrs	No	Dep	CES-D
Schnur 2009	USA	44	NR	0-III	CBT, hypnosis	CBT	Ind	NR	Varied	No	anx, mood/ distress	9 item mood report form, STAI, POM-SV

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures
Schover 2011	USA	300	54	NR	education, peer counselling	Support	Ind	3	1-1.5 hrs	No	Mood/ distress	BSI-18
Simpson 2001	Canada	89	50	0-II	PMR, hypnosis, stress management, CBT	CBT	Grp	6	1.5 hrs	No	Dep, mood/ distress	BDI, POMS
Spiegel 1981	USA	86	55	IV	SET	Support	Grp	>52	1.5 hrs	No	Anx, dep, mood/ distress	POMS
Stanton 2005	USA	558	58	I-II	Education	PsyEd	Ind	2	Varied	No	Dep	CES-D, IES-R
Taylor 2003	USA	90	54	I-IIIa	Education, relaxation, coping strategies, CBT,	CBT	Grp	8	2 hrs	No	Mood/ distress	POMS, MHI, IES
Vos 2007	Holland	87	49	0-III	Experiential-existential therapy	Other	Grp	12 (+ 2 follow-up)	12 hrs	No	Mood/ distress	POMS
Wengstrom 1999	Sweden	134	60	0-IV	Support	Support	Ind	5	30 mins	No	Mood/ distress	IES
Wurtzen 2013	Denmark	366	54	I-III	MBSR	Mindfulness	Grp	8 (+silent retreat)	2 hrs (+ 5 hrs)	No	Anx, dep	CES-D
Yates 2005	Australia	110	49	I-II	Education	PsyEd	Ind	3	10-20 mins	No	Anx, dep, mood/ distress	HADS
Zgaia 2016	Romania	102	59	I-III	Autogenous training, counselling	Support	Ind	1	50 mins	No	Anx, dep	Numerical rating scale

Source	Country	N	Age (mean)	Stage	Treatment components	Treatment type	Treatment format	Treatment sessions	Length of sessions	Exclusively psychologically distressed patients	Outcomes	Outcome measures

Notes. Cpl = couple; Grp = group; Ind = individual CBT = cognitive behavioral therapy; CBSM = cognitive behavioral stress management; MBCR = mindfulness based cancer recovery; MBSR = mindfulness based stress reduction; BATD = behavioral activation treatment for depression; STPP = short-term psychodynamic psychotherapy; SET = supportive expressive therapy; PMR = progressive muscle relaxation; PsyEd = psychoeducation; Anx = anxiety; Dep = depression; mins = minutes; hrs, hours; NR = not reported; HADS = Hospital Anxiety and Depression Scale; CES-D = Center for Epidemiologic Studies Depression Scale; IES = Impact of Events Scale; IES-R = Impact of Events Scale – Revised; POMS = Profile of Mood States; POMS-SF = Profile of Mood States - Short Form; POMS-LASA = Profile of Mood States—Linear Analog Self-Assessment; BDI = Beck Depression Inventory; BAI = Beck Anxiety Inventory; STAI = State-Trait Anxiety Inventory; MHI = Mental Health Inventory; SAS = Self-Rating Anxiety Scale; DASS = Depression, Anxiety, and Stress Scale; HAM-D = Hamilton Depression Rating Scale; SCL-90-R = Symptom Checklist-90-Revised; PAL-C = Profile of Adaptation to Life Clinical Scale; ABS = 40-item Affects Balance Scale; PANAS = Positive and Negative Affect Schedule; HAM-A = Hamilton Anxiety Rating Scale

3.3.1 Aim 1: Overall quality of trials

The mean total quality score on the POMRF was 13.3 out of 38 with median 13 and IQR 6 (i.e. 35% of the maximum possible score; median 34%, IQR 16%). Only 12 trials (13%) reached the criterion of 19, indicating that 79 trials (87%) were of inadequate quality. The mean total quality on the "generic design elements" subscale was 10.5 out of 26 with median 11 and IQR 3 (i.e. 40% of the maximum possible score; median 42%, IQR 12%), while the mean total quality on the "psychotherapy-specific design elements" subscale was 2.8 out of 12 with median 2 and IQR 3 (i.e. 23% of the maximum possible score; median 17%, IQR 25%). In general, therefore, quality was poor, particularly for design elements specific to psychotherapy RCTs.

3.3.2 Aim 2: Quality of specific design elements

Table 3.4 displays the individual item quality scores for each trial. Four generic design elements were particularly poorly implemented: representativeness of the sample (item 3), use of blind evaluators (item 7), control of concomitant treatments (item 18), and evaluation of clinical significance (item 21). These are evaluated in detail below, followed by the elements specific to psychotherapy RCTs.

3.3.2.1 Generic design elements

Representativeness of sample

Of the 74 trials in which emotional distress was the primary outcome, only 11 (15%) specified as an inclusion criterion that patients were distressed. Eight defined distress as scoring above an established cut-off on a specific outcome measure, one defined distress as meeting diagnostic criteria for a depressive disorder, and two defined distress as scoring above an established cut-off on a specific outcome measure and meeting diagnostic criteria for a depressive disorder.

Nine trials excluded patients with a mental health disorder but did not state what disorders these included; four excluded patients with clinical levels of emotional distress; three excluded patients with 'severe' mental health disorders but did not state what disorders these included; two excluded patients with 'prior psychiatric morbidity' but did not elaborate; two excluded patients with a prior history of

psychiatric treatment for a major depressive episode; one excluded patients with a prior history of hospitalisation for a mental health disorder; one excluded patients with current suicidal thoughts; and one excluded patients with previous suicidal thoughts.

Of the 17 trials in which emotional distress was the secondary outcome, only three (18%) specified that to be included, patients must be experiencing the specific difficulty that the primary outcome measured; for example, requiring evidence of insomnia for inclusion in a trial in which the primary outcome was insomnia.

Use of blind evaluators

Only 14 trials (16%) reported using blinded assessors and none employed checks to ensure that the assessor was blind to the treatment condition.

Control of concomitant treatments

Only nine trials (10%) described controlling for concomitant treatments for emotional distress. Of these, two ensured that patients received no additional treatment (psychological or pharmacological); four excluded patients receiving additional psychological treatment (but not those receiving pharmacological treatment); the remaining three included patients taking anxiolytic or antidepressant medication provided the dose was stable (but did not exclude patients receiving additional psychological treatment).

Clinical significance

Only 10 trials (11%) evaluated the clinical significance of treatment effects. In six trials, clinical significance was defined as scoring below a pre-defined cut-off on a specific outcome measure; in three, it was defined as having no diagnosis of anxiety or depression according to DSM criteria; in another three, it was defined as making an arbitrary percentage reduction in score on a specific outcome measure; and in two, it was defined as making a reduction in score on a specific outcome measure based on data-informed reliable change index (RCI). All 10 trials used a different operational definition of clinical significance preventing assessment of the absolute efficacy across different treatments and trials.

3.3.2.2 Psychotherapy-specific design elements

Manualised treatment

Only 46 trials (51%) used a manual to standardise treatment. Of these, 19 referenced a published manual; six referred to an unpublished manual available on request; and 21 referred to a "manual" or a "manualised treatment" without information on how to obtain it. An additional eight trials were ambiguous about whether a manual existed, for example reporting that treatment was *"based on"* or *"modelled after"* a specific manualised treatment. Three more trials reported the treatment was described elsewhere, but the referenced sources provided no information regarding a treatment manual.

Therapist training

Only 29 trials (32%) included therapists qualified to deliver psychotherapy. Of the 29 trials including qualified therapists, only seven (8% of all trials) included therapists with supplementary training in the treatment being investigated.

An additional 29 trials (32%) used therapists with training in the treatment being investigated, but these therapists either had little experience in psychotherapy, for example being "*master's level registered nurse therapists"*, or their clinical background was not reported. The amount of training these therapists received ranged from 6 hours to a ten-week long training course. Of the remaining trials, 14 used therapists with little experience and without training in the treatment being evaluated; in nine it was merely stated that therapists were "trained" or "experienced"; and 10 provided no information about therapist training.

Checks for therapist adherence & competence

Only 17 trials (15%) monitored therapist adherence and only five (4%) monitored therapist competence. An additional 12 reported monitoring treatment delivery but without specifying what aspects were monitored.

Number of therapists

Most trials (67%) included more than one therapist to deliver treatment. However, of these, 16 (18% of all trials) did not specify the number of therapists, although using

the plural “therapists”, and only six (7% of all trials) analysed the effect of therapist on outcome.

Equality of therapy hours

Eleven trials (12%) used only a WLC design. Of those using active controls or TAU, only 19 (21% of all trials) equalised the number of treatment hours between conditions. Of the remaining 61 trials, 26 had more than a 20% difference in treatment hours between conditions, 30 did not report the number of hours received in the control condition, and five did not report the number of hours received by either condition.

3.3.3 Aim 3: Quality comparison with mental health populations

To assess the quality of RCTs in mental health populations relative to RCTs in BCa, eight meta-analyses using the POMRF were identified addressing obsessive compulsive disorder (OCD) in children (Öst et al., 2016) and adults (Öst et al., 2015), anxiety disorders in children (Öst & Ollendick, 2017) and adults (Swain et al., 2013), depression in children (Arnberg & Öst, 2014) and military veterans (Hundt, Barrera, Robinson, & Cully, 2014), and across several mental health populations (e.g. anxiety, depression, borderline personality disorder; Öst, 2014; Sloan et al., 2017). Mean total quality scores ranged from 45 to 56% of the maximum possible, higher than the corresponding score in the present review (35% of the maximum, Table 3.5). Quality of generic and psychotherapy-specific design elements could not be compared because the mental health meta-analyses did not report individual item scores.

3.3.4 Aim 4: Quality comparison by treatment type

One-way ANOVAs revealed no significant difference in overall quality scores or generic subscale scores between different treatment types. A significant difference was seen for psychotherapy-specific subscale scores (Table 3.6). Post-hoc testing showed that MBT trials were of better quality than support, peer and “other” treatment trials. However, MBT trials still only had a mean quality score of 5.4 out of 12 (i.e. 45% of the maximum possible score) on this subscale.

3.3.5 Aim 5: Quality trends over time

The overall quality score modestly improved with year of publication (rho = 0.4, $p < 0.01$; Figure 3.2). However, the mean total quality of the 30 trials published in the last five years was still only 14.7 with median 14.5 and IQR 7 (i.e. 38% of the maximum possible, median 34%, IQR 13%) and only six of these met our criterion for adequate quality. Generic design elements improved across publication year (rho = 0.48, $p < 0.01$) but psychotherapy-specific design elements did not (rho = 0.15, $p = 0.15$; Figure 3.3).

Table 3.4: POMRF scores for each included trial

Author	#1	#3	#5	#6	#7	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	Total quality score
Abad 2016	1	0	2	2	0	1	2	0	1	0	0	0	0	0	1	0	1	0	2	**13**
Allard 2007	1	0	1	2	0	1	0	1	1	0	0	0	0	0	0	0	2	0	0	**9**
Allen 2002	2	0	1	1	0	1	0	0	1	0	1	0	0	0	0	1	1	0	0	**9**
Andersen 2004	2	0	2	2	0	1	0	1	2	0	1	1	1	0	0	1	1	0	0	**15**
Anton 2015	1	1	2	0	0	1	2	0	1	0	0	0	0	0	0	0	1	0	0	**9**
Antoni 2001	1	0	2	2	0	1	1	0	2	2	1	0	0	0	0	1	1	0	0	**14**
Antoni 2006	2	0	2	1	0	1	1	0	1	2	1	0	0	0	0	2	1	0	0	**14**
Antoni 2009	2	0	2	2	0	1	1	0	1	2	1	0	1	0	0	1	1	0	0	**15**
Arving 2007	2	0	2	0	0	1	2	1	1	0	2	0	0	0	0	2	2	1	0	**16**
Ashing 2014	2	2	2	2	0	1	0	0	0	0	1	0	1	0	0	0	1	1	0	**13**
Badger 2007	1	0	2	2	0	1	1	0	1	0	1	1	0	0	0	0	2	0	0	**12**
Badger 2013a	1	0	2	2	0	1	2	0	1	0	0	0	0	0	0	1	2	0	2	**14**
Badger 2013b	2	0	2	1	0	1	2	0	1	0	0	0	2	0	0	1	2	0	2	**16**
Baucom 2009	1	0	1	2	0	1	0	0	2	0	1	0	0	0	0	0	1	0	0	**9**

Author	#1	#3	#5	#6	#7	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	Total quality score
Beutel 2014	1	1	2	1	1	1	0	2	1	2	1	1	0	0	1	2	2	1	0	**20**
Bjornklett 2012	1	0	2	1	0	1	0	2	1	0	1	0	0	0	0	1	1	0	0	**11**
Boesen 2011	1	0	2	1	0	1	0	0	2	0	1	1	0	0	0	1	1	0	0	**11**
Bower 2015	1	0	2	0	0	1	0	2	1	0	0	0	0	0	0	1	2	0	2	**12**
Bredal 2014	1	0	2	1	0	1	2	0	1	1	1	0	0	0	0	0	2	0	0	**12**
Budin 2008	1	0	1	2	0	1	1	2	1	2	1	0	0	0	0	1	1	0	0	**14**
Burton 1995	0	0	2	1	0	1	0	0	2	0	0	0	0	0	0	0	1	0	0	**7**
Carlson 2013	1	2	2	0	1	1	2	1	0	2	1	1	0	0	0	1	2	0	2	**19**
Chan 2006	1	0	2	0	0	1	2	0	1	2	0	0	0	0	0	0	1	0	2	**12**
Christensen 1983	1	0	2	0	0	1	0	0	0	0	0	0	1	0	0	2	0	0	0	**7**
Classen 2001	2	0	2	2	0	1	0	0	2	0	1	0	0	0	0	1	1	0	0	**12**
Classen 2008	2	0	2	1	0	1	0	2	2	2	2	0	0	0	0	0	2	0	0	**16**
Cohen 2007	1	0	1	2	1	1	2	0	1	0	0	0	0	0	0	1	2	0	2	**14**
Cousson-Gelie 2011	1	0	2	2	0	1	2	1	0	0	1	1	0	0	0	0	2	0	2	**15**
Creuss 2000	1	0	2	2	0	1	0	0	0	0	1	0	1	0	0	0	2	0	2	**12**
Davis 1986	1	0	0	2	0	1	2	0	1	0	0	1	0	0	0	1	1	0	2	**12**

Author	#1	#3	#5	#6	#7	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	Total quality score
Dirksen 2007	2	1	2	2	0	1	1	2	0	0	0	0	0	0	1	0	2	0	2	**16**
Dolbeault 2009	1	0	2	1	0	1	0	0	1	2	1	1	0	0	0	1	2	0	2	**15**
Dowlatabadi 2016	1	1	2	2	0	1	0	0	0	0	0	0	0	0	1	0	2	0	0	**10**
Edelman 1999a	1	0	1	1	0	2	2	0	1	0	1	1	0	0	1	0	2	0	2	**15**
Edelman 1999b	1	0	2	1	0	1	0	0	1	0	1	1	0	0	0	0	1	0	0	**9**
Edmonds 2001	1	0	2	1	0	1	0	0	1	0	1	2	0	0	0	0	1	0	0	**10**
Ferguson 2012	1	1	2	1	1	1	0	2	1	0	0	0	0	0	0	2	1	0	2	**15**
Ferguson 2016	1	2	2	0	1	1	1	2	1	2	0	1	1	0	0	0	2	0	2	**19**
Fillion 2008	1	0	1	2	0	1	0	2	1	0	1	0	2	0	0	0	2	0	0	**13**
Fukui 2000	2	0	2	1	0	1	0	0	1	0	1	1	0	0	0	1	2	0	2	**14**
Gaston 2000	1	0	2	1	0	1	0	1	0	0	0	0	0	0	0	0	2	0	0	**8**
Giese-Davis 2016	2	0	2	2	0	1	0	1	1	2	1	1	0	0	0	2	1	0	0	**16**
Goodwin 2001	1	0	2	0	0	1	0	1	0	2	1	2	0	0	0	1	2	0	0	**13**
Gotay 2007	1	0	1	1	0	1	0	1	1	0	1	0	0	0	0	0	1	0	0	**8**
Graves 2003	1	0	2	1	0	1	0	0	0	2	1	0	0	0	0	1	1	0	0	**10**
Groarke 2012	2	0	2	2	0	1	0	1	2	1	0	1	0	0	0	1	2	0	0	**15**

Author	#1	#3	#5	#6	#7	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	Total quality score
Gudenkauf 2015	2	0	1	2	1	1	1	2	0	0	1	0	1	1	0	0	2	0	2	**17**
Heiney 2003	1	0	2	2	0	1	0	0	1	0	1	1	0	0	0	2	2	0	0	**13**
Henderson 2012	2	0	2	1	0	1	1	2	2	2	1	1	0	0	0	1	1	0	2	**19**
Hoffman 2012	1	0	2	1	1	1	0	2	1	1	0	2	0	0	0	1	2	0	2	**17**
Hopko 2011	2	2	2	2	0	2	2	2	2	2	1	0	2	2	1	1	2	1	2	**30**
Kalaitzi 2007	0	0	2	1	0	1	0	0	0	0	0	0	0	0	0	2	1	0	0	**7**
Kissane 2003	1	0	2	0	0	2	0	1	1	0	2	0	2	0	0	1	1	0	0	**13**
Kissane 2007	2	0	2	0	0	1	0	2	2	2	1	1	2	2	0	1	1	1	0	**20**
Lane 2005	1	0	2	1	0	1	0	0	1	0	0	1	0	0	0	2	2	0	2	**13**
Larson 2000	0	0	2	0	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	**6**
Lechner 2014	2	1	2	2	1	1	1	0	1	2	0	1	0	0	0	1	2	0	2	**19**
Lee 2013	1	2	2	2	0	1	0	1	0	0	1	0	0	0	0	2	1	0	0	**13**
Lengacher 2009	2	0	2	0	0	1	0	2	0	1	0	2	1	1	0	2	1	0	2	**17**
Lewis 2015	1	0	1	2	1	1	1	1	1	0	1	0	0	0	0	2	2	0	0	**14**
Maguire 1980	0	0	2	0	0	1	0	0	2	0	0	0	0	0	0	0	0	1	0	**6**
Manne 2005	2	0	2	2	0	1	0	0	1	2	2	0	2	0	0	2	2	0	0	**18**

Author	#1	#3	#5	#6	#7	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	Total quality score
Manne 2016	2	0	1	2	0	1	2	0	2	2	1	0	1	0	0	1	2	0	2	**19**
Marchioro 1996	1	0	2	0	0	1	0	0	1	0	0	1	0	0	0	0	2	0	0	**8**
Marcus 2010	2	0	1	1	0	1	0	0	1	0	1	0	0	0	0	0	1	0	0	**8**
McArdle 1996	1	0	1	0	0	1	2	0	2	0	0	0	0	0	0	1	2	0	0	**10**
Mens 2015	1	0	2	2	0	1	1	0	1	0	1	0	1	0	0	2	1	0	0	**13**
Miyashita 2005	2	0	2	1	0	1	0	0	1	0	0	0	0	0	0	0	2	0	0	**9**
Napoles 2015	2	0	1	2	1	1	0	0	1	0	1	0	0	0	0	1	2	0	0	**12**
Naumann 2012	1	0	2	0	0	1	2	0	0	0	1	1	0	0	0	1	1	0	0	**10**
Pelekasis 2016	1	0	2	2	0	1	1	0	0	0	1	0	0	0	2	0	1	0	0	**11**
Qiu 2013	2	2	2	1	1	1	0	0	1	0	0	1	1	1	2	1	2	0	2	**20**
Rissanen 2015	1	1	2	1	0	1	2	2	1	0	1	0	0	0	0	1	2	0	0	**15**
Samarel 1997	1	0	1	1	0	0	2	0	1	0	2	0	0	0	0	1	2	0	2	**13**
Sandgren 2000	1	0	2	2	0	1	0	0	1	0	1	0	0	0	0	0	2	0	0	**10**
Sandgren 2003	1	0	2	2	0	1	0	0	1	0	1	0	0	0	0	0	1	0	0	**9**
Savard 2005	2	1	2	1	0	1	0	2	2	0	0	1	0	0	1	1	2	1	2	**19**
Savard 2006	2	2	2	0	1	1	0	0	1	1	1	2	0	0	1	1	2	1	2	**20**

Author	#1	#3	#5	#6	#7	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	Total quality score
Savard 2014	1	2	2	1	0	1	1	2	2	0	1	1	1	1	0	2	2	1	0	**21**
Scheier 2005	2	0	2	2	0	1	1	0	1	0	1	0	1	0	0	1	2	0	0	**14**
Schnur 2009	1	0	2	2	1	1	0	2	1	0	1	1	0	0	0	0	1	0	0	**13**
Schover 2011	1	0	1	0	0	1	1	0	2	0	2	0	0	0	0	0	0	0	0	**8**
Simpson 2001	1	0	2	0	0	1	0	0	2	0	0	1	0	0	0	0	2	0	0	**9**
Spiegel 1981	0	0	2	0	0	1	0	0	1	0	1	0	0	0	0	0	1	0	0	**6**
Stanton 2005	1	0	2	1	0	1	1	2	1	0	1	0	1	1	0	2	1	0	0	**15**
Taylor 2003	2	0	2	2	0	1	0	0	2	2	1	0	0	0	0	1	1	0	0	**14**
Vos 2007	1	0	1	2	0	1	2	0	2	0	1	2	0	0	0	1	1	0	2	**16**
Wengstrom 1999	1	0	2	2	0	1	1	0	1	0	1	0	0	0	0	2	2	0	0	**13**
Wurtzen 2013	1	0	2	1	0	1	0	1	1	2	1	2	2	0	0	2	1	1	0	**18**
Yates 2005	1	0	2	0	1	1	1	1	1	0	1	0	0	0	0	1	1	0	2	**13**
Zgaia 2016	1	0	1	0	1	0	0	1	1	0	0	1	0	0	0	2	2	0	0	**10**

Table 3.5: Comparison of total quality scores on the POMRF of RCTs in this population with RCTs in mental health populations

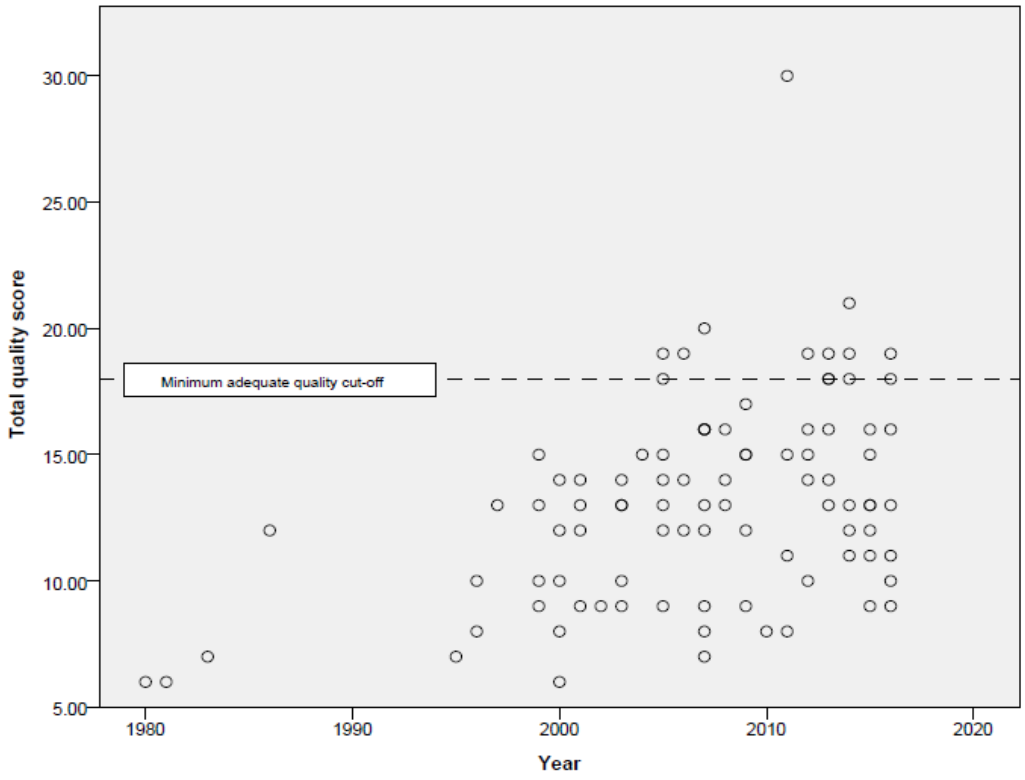
	Number of RCTs included	Mean total score	Maximum possible score	Maximum possible score as a %	Range of scores	Range of scores as a %
Anxiety disorders in children	23	21.6	44	49%	-	-
Anxiety disorders in adults	4	20	44	45%	14-27	32-61%
Depression in children	10	22	44	50%	8-30	18-69%
Depression in military veterans	5	24	44	55%	15-32	34-63%
Emotional distress in BCa	91	13.3	38	35%	6-30	16-79%
Individuals with anxiety, depressive, or borderline personality disorders	34	24	44	55%	14-34	32-77%
Individuals with mental health conditions	31	20.3	44	46%	-	-
OCD in adults	37	23	44	52%	15-34	34-77%
OCD in children	25	24.6	44	56%	-	-

Note. POMRF = psychotherapy outcome study methodology rating form; RCT = randomised controlled trial; BCa = breast cancer; OCD = obsessive compulsive disorder

Table 3.6: Means (SDs) and F-values for the items on the POMRF for different treatment types

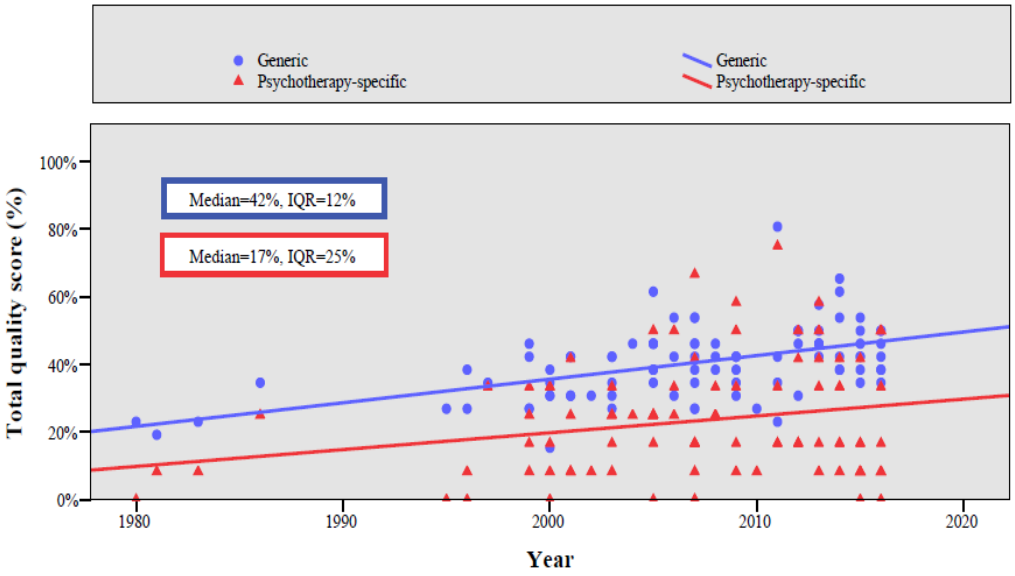
	CBT (n=40)	Mindfulness (n=5)	PsyEd (n=6)	Support (n=21)	Other (n=13)	Peer-led (n=6)	All trials (n=91)	F-value
Total quality score	14.13 (4.61)	16.6 (2.7)	13.17 (0.75)	12.14 (3.86)	12 (3.96)	11.66 (3.14)	13.27	1.78
Subscale 1: Generic design elements	11.08 (3.2)	11.2 (1.3)	11 (1.26)	9.67 (2.31)	9.85 (2.76)	9.83 (2.64)	10.49	1.05
Subscale 2: Psychotherapy-specific design elements	3.05 (1.99)	5.4 (2.07)[a]	2.17 (0.75)	2.48 (2.18)[b]	2.15 (1.77)[b]	1.83 (1.17)[b]	2.78	2.81*

Note. POMRF = psychotherapy outcome study methodology rating; PsyEd = psychoeducation; *P<0.01; [a,b]Means with different superscript differ significantly



Abbreviations: POMRF, psychotherapy outcome study methodology rating form.

Figure 3.2: Scatterplot of total quality scores on the POMRF by year of publication



Abbreviations: POMRF, psychotherapy outcome study methodology rating form.

Figure 3.3: Scatterplot of POMRF scores on generic design elements and psychotherapy-specific design elements by year of publication

3.4 Discussion

Findings from this study show that the methodological quality of RCTs for emotional distress in BCa is poor. Most RCTs were of inadequate quality and of lower quality than those in mental health. While quality modestly improved from 1980 to 2016, most of the more recently published RCTs were still poorly designed.

Quality was particularly poor for design elements specific to psychotherapy RCTs. Moreover, while implementation of generic design elements improved over time, that of psychotherapy-specific ones did not. Design elements specific to psychotherapy RCTs were lacking in most trials, thereby compromising the internal validity of such trials. Only around one in 20 monitored therapist competence or analysed the effect of therapist on outcome; only around one in 10 employed therapists who were adequately trained in the treatments or monitored therapists' adherence to them; only around one in five compared conditions with an equal number of treatment hours; and barely half used a manual to standardise treatment. The purpose of these design elements is to ensure that a psychological treatment is implemented correctly. Because none of the

trials in this review adequately implemented all these elements, their findings provide unreliable information about the treatments the authors are claiming to evaluate.

Although, in general, generic design elements were not as poorly conducted as those specific to psychotherapy RCTs, important ones were still neglected, thereby further compromising trials' internal and external validity. Only around one in 10 trials controlled for concomitant treatments, blinded assessors, specified as an inclusion criterion that participants were distressed, or evaluated the clinical significance of treatment effects.

While all these areas represent significant deficits in the BCa literature, the latter two are particularly concerning. Consensus-based clinical practice guidelines specifically recommend the use of psychological treatments for BCa patients experiencing clinical levels of emotional distress (Coleman, Hession, & Connolly, 2011; Holland, 1999) and the standards by which a treatment is considered "evidence-based" in the United States recommend that trials provide an estimation of clinical significance (Chambless & Hollon, 1998; Tolin, McKay, Forman, Klonsky, & Thombs, 2015). If trials do not target patients with emotional distress, findings cannot be generalised to the population of patients to whom the treatments would be offered in practice: i.e. those with clinical levels of emotional distress. Additionally, if trialists do not report the clinical significance of treatment effects, it is difficult for researchers, clinicians, service providers and policy-makers to assess the practical relevance of findings. Determining and applying standardised criteria for the clinical significance of treatment effects would advance psychotherapy outcome research in BCa. The most established method for determining clinical significance is the approach developed by Jacobson and colleagues (Jacobson, Follette, & Revenstorf, 1984; Jacobson & Truax, 1991; see chapter 4 for more details).

The overall methodological quality of RCTs did not differ by type of treatment. For design elements specific to psychotherapy RCTs, quality did differ by type of treatment explored, with MBT trials achieving better quality scores than support, peer, and "other" treatment trials However, trials evaluating MBT were still of limited quality. Many of the trials provided insufficient details about treatment methods and procedures. Thus, it was not possible to categorise treatments by distinct treatment approaches, but only by broad categories of type of treatment. This highlights the

importance of studies clearly reporting the type of treatment used with clear and unambiguous descriptions of the treatments being compared.

Although the methodological quality of RCTs modestly improved over time, the majority of the most recent RCTs were still of inadequate quality. Thus, there is clearly substantial room for improvement in the conduct of RCTs in this population.

This review has some important limitations. The assessment of trial quality was limited to what was included in the published reports. As most journals impose word limits, authors may have excluded important information. Thus, some trials may have implemented unreported design elements. This review also relied on summary scores to quantify the overall quality of trials. Summary scores can be problematic as they can mask methodological strengths and weaknesses of a trial. Trials that differ in the conduct of individual design elements may still result in the same overall score. Finally, whether methodological quality differed amongst trials including patients at different points in the BCa trajectory was not evaluated because many trials included patients at multiple points in the disease trajectory.

In conclusion, the current view that efficacious psychological treatment exists for emotional distress in BCa patient**s** is based on poor quality RCTs. It does not follow that efficacious treatments do not exist, or that conclusions of previous meta-analyses are wrong. However, with increasing investment in, and growing priority of, psychological treatments for emotional distress in BCa (Hewitt, Herdman, & Holland, 2004; Jassim et al., 2015) it is imperative that future psychotherapy RCTs are conducted with greater methodological rigour to make sure evidence-based practice occurs in clinical settings. Researchers need to ensure that the methodological issues presented in this review are adequately implemented in future trials. Trials need to include participants with clinical levels of emotional distress and report the clinical significance of treatment effects. Trialists must also consider the methodological challenges specific to psychotherapy RCTs, particularly manualisation.

If relevant health policies are to be adequately empirically informed, meta-analyses must also account for important methodological issues presented in this review. However, no previous meta-analysis has focused specifically on clinically distressed patients, excluded non-manualised treatments or examined whether manualisation influenced treatment efficacy, or examined whether treatment effects are clinically

significant. Therefore, the practical relevance of previous meta-analysis is questionable. In chapter 5, a meta-analysis accounting for these methodological limitations is presented. First, however, the development of clinical significance is outlined, and a review and critique of a specific approach for evaluating clinical significance, known as the 'Jacobson method' is presented in chapter 4.

Chapter 4. Measuring Clinical Significance in Psychotherapy Outcome Research: A Review of Clinical Significance Methodology

4.1 Introduction

Most psychotherapy outcome studies aim to demonstrate the superior efficacy of one treatment over alternative treatments or control conditions. This has typically been conveyed by comparing group means using inferential statistics (Ogles, Lunnen, & Bonesteel, 2001). Traditionally, effect sizes are used to express whether the average amount of change in one group is superior to that in another. An effect is deemed *statistically significant* if the magnitude of the mean difference is beyond the range of chance (conventionally set at $p<0.05$). The evaluation of effect sizes in psychotherapy outcome research has long been advocated (Cohen, 1995; Thompson, 2002); and the fifth edition of the APA publication manual stresses the importance of reporting effect sizes*: "For the reader to fully understand the importance of your findings, it is almost always necessary to include some index of effect size"* (APA Publication Manual, 2001 pp. 25-26). Although effect sizes and associated tests of significance provide valuable group information, they provide no information about individual variability in treatment response. Thus, it is possible to have a large effect size in favour of the treatment group but still have many or even the majority of treated patients present with substantial and impairing symptoms. Indeed, the size of an effect can often have no bearing on the clinical importance of a treatment. Jacobson and Truax (1991) used the following example to demonstrate how a large effect size can have no clinical importance:

> *"If a treatment for obesity results in a mean weight loss of 2 lb and if subjects in a control group average zero weight loss, the effect size could be quite large if variability within the groups were low. Yet the large effect size would not render the results any less trivial from a clinical standpoint"* (p. 12).

Thus, while necessary, it is not enough to show that a treatment performs better than a control condition or another active treatment at the group level. An evaluation of clinical significance is needed to indicate the proportion of patients who benefit from treatment.

Clinical significance has been operationalised in numerous ways, all with the aim of determining if a treatment results in a clinically meaningful change. The most established and widely used method for determining clinical significance in psychotherapy outcome research is the approach developed by Jacobson and

colleagues (Jacobson, Follette, & Revenstorf, 1984; Jacobson and Truax, 1991), referred to as the 'Jacobson method'. This chapter provides a review of clinical significance from its development in behaviour analysis to the now widely used Jacobson method. Strengths and weakness of the Jacobson method and alternative methods are also considered.

4.2 Development of clinical significance

The current concept of clinical significance in psychotherapy outcome research stems from Risley's seminal paper (1970) in the field of behaviour analysis. He proposed that treatment outcomes should be assessed according to both experimental and therapeutic criteria. The experimental criterion attempts to determine whether a treatment is responsible for behaviour change (i.e. by comparing behaviour during treatment with what it would be like without treatment); while the therapeutic criterion attempts to determine whether changes in behaviour are meaningful to the individual. Defining this therapeutic or 'clinically significant' criterion proved difficult. In some cases, the therapeutic criterion can be easily applied. For example, the elimination of seizures in a person with epilepsy would be universally accepted as clinically significant. However, a 30% reduction in seizures would likely represent a meaningful improvement but may not represent a clinically significant one. Thus, when treatment success is measured by the intensity or frequency of a behaviour (i.e. severity of emotional distress) defining what is clinically significant is much more difficult.

In the 1970s, social validity (Kazdin, 1977; Wolf, 1978) emerged as a potential solution to this problem. Social validity was founded in applied behavioural analysis and refers to the social importance of treatment. Several facets of social importance can be distinguished but the focus here, and within the broader field of psychotherapy outcome research, is behaviour change. According to the social validity approach, what is deemed as clinically significant can be assessed in two specific ways: 1) subjective evaluation or 2) social comparison with well-functioning peers.

The 'subjective evaluation' approach involves asking those in everyday contact with the treated patient to evaluate the treated patient's behaviour. This enables the researcher to see if the patient has made changes that are observable by others. Strupp and Hadley (1977) extended the original method of subjective evaluation by

suggesting three specific perspectives be considered: 1) the patient themselves, 2) the therapist, and 3) those in everyday contact with the patient.

The 'social comparison' approach involves a social comparison with a well-functioning population. The underlying premise is that if a treatment has resulted in a clinically significant change then a patient's post-treatment behaviour should be indistinguishable from that of well-functioning peers (Kazdin, 1977). This approach led to an increased recognition of the need to examine the clinical relevance of changes occurring during psychotherapy. The emerging consensus was that it would be useful to objectify what constitutes a return to functionality (Kazdin & Wilson, 1978). Initial definitions were either arbitrary (i.e. a 50% reduction in symptoms; Jansson & Öst, 1982) or subjective (i.e. an unspecified level of functioning deemed clinically meaningful without a theoretical basis; Barlow & Mavissakalian, 1981). An approach largely free from bias that could be applied to a wide range of psychological disorders and enable cross-study comparisons was therefore needed to overcome these heterogeneous and methodologically flawed definitions. Jacobson, Follette, and Revenstorf (1986) recognised this and presented a standardised method of defining what constitutes a return to functionality, now coined 'clinical significance'.

4.3 The Jacobson method of clinical significance

Jacobson and colleagues (1984, 1991) believed that clinical significance should meet the standards of efficacy set by consumers (i.e. patients), clinicians and researchers. As patients, clinicians, and researchers often expect psychotherapy to remove the problem that patients bring to therapy, the central premise of the Jacobson method is that a clinically significant change should return patients to normal levels of functioning: "*change in therapy is clinically significant when the client moves from the dysfunctional to the functional range during the course of therapy"* (Jacobson et al., 1984, p. 340). The Jacobson method has two criteria. The first involves the calculation of a cut-off point on a well-validated outcome measure to determine if a patient's posttreatment score has a greater probability of being drawn from a functional or dysfunctional population. Second, the 'reliable change index' (RCI) determines if the extent of change from pre- to post-treatment is statistically reliable. This ensures that the pre- to post-treatment change score is not merely an artefact of measurement error. Applying these two criteria, patients can be classed into one of

four categories: i) ‘recovered’, if they make a statistically reliable change and move from a dysfunctional to a functional population; ii) ‘improved’, if they make a statistically reliable change but do not move from a dysfunctional to a functional population; iii) ‘unchanged’, if they do not make a statistically reliable change; and iv) ‘deteriorated’, if they make a statistically reliable change for the worse.

Criterion 1: Cut-off point. There are three available methods for determining if a patient’s level of functioning is within the functional range following treatment: (a) patient’s post-treatment score falls outside the range of dysfunctionality, defined as falling two or more standard deviations (SDs) beyond the mean of the dysfunctional population, in the direction of functionality; (b) patient’s post-treatment score is within the range of functionality, defined as falling within two SDs of the mean of the functional population; and (c) patient’s post-treatment score has a greater probability of being drawn from the functional than the dysfunctional population.

Based on hypothetical data, Jacobson and Truax (1991) clarified when each cut-off point is most suitable. When functional and dysfunctional distributions are overlapping, cut-off point (a) is the most stringent, cut-off point (b) is the most lenient, and cut-off point (c) falls in the middle of the two cut-offs. When functional and dysfunctional distributions are not overlapping, cut-off point (a) is the most lenient, cut-off point (b) is the most stringent and cut-off point (c) still falls in the middle of the two cut-offs. To calculate these cut-off points, normative data is required. Ideally normative data should be available for a functional as well as a dysfunctional population because if each study uses its own dysfunctional sample then each study will have a different cut-off point, making comparisons between studies difficult (Hollon & Flick, 1988; Wampold & Jenson, 1986). When appropriate normative data exists for a functional and dysfunctional population, cut-off point (c) is the method of choice. As it includes normative data from both functional and dysfunctional populations, it is the least arbitrary method. If normative functional data does not exist, only cut-off point (a) can be used; and if normative dysfunctional data does not exist, only cut-off point (b) can be used.

Criterion 2: Reliable change index. Post-treatment scores may cross the cut-off point for functionality but the change in score may not be statistically reliable. For example, if the cut-off point on an outcome measure is 23 and a patient’s pre-treatment score

falls just outside this cut-off (i.e. a score of 24) but their post-treatment score falls just within (i.e. a score of 22) they would have crossed the cut-off point and be part of the functional population. However, a change score of two is unlikely to be statistically reliable. Thus, Jacobson et al. (1984) proposed an RCI which identifies the threshold beyond which symptoms must change on an outcome measure for it to be considered statistically reliable. Due to mathematical problems identified by Christensen and Mendoza (1986), the original formula was modified by Jacobson and Truax (1991). The modified RCI is calculated by dividing the difference between the post-treatment (xpost) and pre-treatment (xpre) scores by the standard error of differences (Sdiff):

$$RCI = \frac{(\text{Xpost} - \text{xpre})}{Sdiff}$$

The Sdiff can be calculated directly from the standard error of measurement (SE):

$$Sdiff = \sqrt{2(SE)^2}$$

The size of the SE hinges on the reliability (r) of the outcome measure.

$$SE = SD\sqrt{1} - r$$

Thus, the more reliable a measure, the smaller the amount of change needed between pre- and post-treatment scores to achieve a statistically reliable change. While the type of reliability test was not specified by Jacobson et al. (1984, 1991), the use of internal consistency rather than test-retest reliability has since been recommended (Martinovich, Saunders, & Howard, 1996; Tingey, Lambert, Burlingame, & Hansen, 1996a). This is because test-retest scores may reflect changes in score due to treatment opposed to measurement error. An RCI equal to or larger than ±1.96 is required for the change to be statistically reliable at $p<0.05$:

$$RCI = \frac{(\text{Xpost}-\text{xpre})}{Sdiff} * 1.96$$

4.4 Criticisms of the Jacobson method and proposed alternatives

The Jacobson method has been criticised by several authors and in an attempt to provide more precise estimations of clinically significant change, alternative methods have been proposed. The most common criticism of the Jacobson method is not

considering regression towards the mean. Regression towards the mean is the phenomenon that extreme scores on the first measurement will likely be closer to the mean on the second measurement. Thus, patients with more extreme scores pre-treatment (e.g. those with extreme anxiety or depression scores) are more likely to make a statistically reliable change by post-treatment.

Alternative methods have been developed to try and deal with the issue of regression towards the mean. Hsu (1989) developed the Gullisken-Lord-Novick (GLN) method, which modifies the RCI formula to include estimates of the population mean which scores are expected to regress towards. However, population means and SDs are rarely available. Speer (1992) developed the Edwards-Nunnally (EN) method which places confidence intervals (CIs) around estimates of patients' true pre-treatment scores and then calculates their post-treatment score based on this CI. Nunnally and Kotsch (1983) developed the Nunnally-Kotsch (NK) method which estimates patient's true pre-treatment scores based on their observed pre-treatment scores. Lastly, Hageman and Arrindell (1999) developed the Hageman and Arrindell (HA) method which attempts to correct for regression towards the mean by modifying both the RCI and cut-off criteria. Similar to the EN method, the RCI is determined by including both pre- and post-treatment reliability in its calculation; while the calculation of the cut-off point incorporates true score mean equivalents and reliability coefficients to account for measurement error.

The Jacobson method has also been criticised for being too conservative. A severely dysfunctional patient could make a large amount of improvement but still fall within the dysfunctional population. Although such patients would be classed as 'improved', they would not be classed as having made a clinically meaningful change (i.e. 'recovered'). Tingey, Lambert, Burlingame, and Hansen (1996b) argue that a return to the functional population is not always a realistic goal. For example, it is impossible for physical therapy to cure Parkinson's disease (PD). Thus, patients with PD will never be part of the functional population. Yet physical therapy may still facilitate numerous clinically meaningful changes for patients with PD. Similarly, patients with severe psychiatric disorders may never realistically fall within a functional population. Yet, most would agree it would be incorrect to assume that such patients can never make a clinically meaningful change. To overcome this problem, Tingey et al. (1996b) proposed using multiple samples on a continuum. They identified four separate

populations which patients could be classed: asymptomatic, normal, mild disturbance, and severe disturbance. Clinical significance requires movement into the adjacent population in the positive direction. Although this method would provide a more thorough understanding of individual change, this greater delineation in samples would require substantially more normative data, which is already sparse. Jacobson and colleagues (1999) did acknowledge the limitation of their method in not capturing all aspects of clinically meaningful change. However, they argued that the original purpose of operationalising clinically significant change was to meet the standards of efficacy set by consumers (i.e. patients), clinicians and researchers i.e. to remove the problem that patients brought to therapy (Martinovich et al., 1996). Nonetheless, this is still a valid criticism of the method.

Another criticism of the Jacobson method, and the alternatives described above, is their sole reliance on change scores at two time points (i.e. pre- and post-treatment). Speer (1995) advocated for a multi-wave data approach to clinical significance using hierarchical linear modelling (HLM). The HLM method uses longitudinal data and growth-curve analysis to reflect the change that occurs from pre- to post-treatment more precisely.

Several other methodological issues of the Jacobson method exist. First, when defining what constitutes a return to functionality, it is preferable to use normative data on a functional sample. However, defining what constitutes a functional sample is difficult. Jacobson and Revenstorf (1988) stated that while a functional sample should preferably not include patients who are dysfunctional, it should not exclude 'outliers' of dysfunctional patients not receiving therapy. Likewise, Hollon and Flick (1988) proposed that representative functional samples should be based on an unscreened demographically normative population. However, it has been estimated that around 20% of the general population have clinical levels of psychopathology (Saunders, Howard, & Newman, 1988). This poses problems when suggesting that a clinically significant change signifies return to normal levels of functioning as such groups are likely to include a substantial proportion of dysfunctional individuals. Estimates based on unscreened populations will therefore provide a much more lenient evaluation of what constitutes a return to functionality. Thus, it is important that studies clearly define their functional sample: "*it is essential that an investigator's perspective on what is normal be clearly articulated"* (Saunders et al., 1988, p. 210).

Second, the Jacobson method assumes normal distributions. However, many of the outcome measures used in psychotherapy outcome research have a limited range leading to skewed distributions. Skewed distributions are more likely to cause problems when psychopathology measures are used on patients within the functional range (Seggar, Lambert, & Hansen, 2002). Both Tingey et al. (1996b) and Jacobson and Revenstorf (1988) recognised the potential problem of non-normal distributions but no solution has yet been provided. However, this is more of a limitation of available outcome measures opposed to the Jacobson method itself.

Third, the Jacobson method is unable to adequately categorise change in patients who enter treatment already in the functional range (Lambert & Ogles, 2009). As these patients cannot move from a dysfunctional to a functional population, it is impossible for them to be classed as 'recovered'. This is rarely the case as patients' who are part of the functional population prior to treatment usually do not require such treatment. However, it is still possible to apply Jacobson's second criterion (RCI) to these patients and evaluate whether they have made statistically reliable change. Therefore, these patients can still be classed as 'improved', 'unchanged, or 'deteriorated'.

Finally, one of the main aims of developing the Jacobson method was to enable meaningful comparisons across studies. However, the Jacobson method can be applied to any outcome measure. Different outcome measures will obviously result in different cut-offs and RCIs. Therefore, the proportion of patients classed as having achieved clinically significant change will differ depending on the outcome measure used. For example, Ogles, Lambert, and Sawyer (1995) evaluated the clinical significance of CBT for patients with depression using the Jacobson method. They found that different measures produced differing results. At post-treatment, 28% of patients recovered according to the Beck Depression Inventory (BDI), while 45% recovered according to the Hamilton depression scale (HAM-D). This highlights the need to develop core outcome measures to be used within trials in a specific area.

4.5 Convergences and divergences in classification between different methods

Following the emergence of alternative clinical significance methods, it is important to consider whether different methods yield different results. Six empirical studies have investigated the convergences and divergences in classification between methods (Lambert & Ogles, 2009). Their findings will be discussed below.

Speer and Greenbaum (1995) compared rates of meaningful change amongst 73 patients who were diagnosed with a range of psychological disorders based on five clinical significance methods: the Jacobson (RCI criterion only), EN, NK, GLN, and HLM methods. High rates of agreement in classification (78-81%) were found between methods, apart from the GLN method (51%). Speer and Greenbaum (1995) therefore recommended the use of the Jacobson method as it evades statistical issues associated with residualised true score adjustments, it is the most straight-forward method, and it is the most widely used method allowing comparisons across studies. They also concluded that regression towards the mean is not nearly as big of a problem as previously thought.

McGlinchey and Jacobson (1999) compared the Jacobson and HA methods across 30 couples receiving couple's therapy. No substantial differences between methods emerged. As the HA method did not yield a more sensitive estimate of clinically significant change (the main reason for its development), McGlinchey and Jacobson (1999) concluded that the Jacobson method is preferable.

McGlinchey, Atkins, and Jacobson (2002) compared rates of change amongst 129 patients receiving psychological treatment for depression. The same five methods as Speer and Greenbaum (1995) were compared, apart from the NK method which was replaced by the HA method. No significant differences in classifications emerged between the different methods. As the methods were essentially equivalent, McGlinchey et al. (2002) supported the use of the Jacobson method as *"it has yet to be rejected by an alternative method of superior performance"* (p542).

Bauer, Lambert, and Nielsen (2004) compared rates of change across the same five methods as McGlinchey et al. (2002) amongst 386 patients receiving psychological treatment across a range of psychological disorders. The average amount of agreement in classification between methods ranged from 71-85%. The HA method produced the most cautious recovery rates (12%) and the EN method the most lenient (21%). Due to its simplicity and because it provides a reasonable position between the extremes of the HA and EN methods, the Jacobson method was recommended.

Atkins, Bedics, McGlinchey, and Beauchaine (2005) compared four methods: the Jacobson, GLN, EN, and HA methods. Unlike previous comparison studies, Atkins et al. (2005) used simulated data to explore differences between methods. Overall, there

was considerable agreement in classification between methods, especially when the reliability of the measure was high. The authors concluded that, as no one method could be preferred over another for statistical reasons, the Jacobson method is preferable because of its wide use and ease of computation.

Ronk, Hooke, and Page (2012) compared five methods (the Jacobson, NK, EN, HA and GLN methods) across a wide range of outcome measures in inpatients (N=2,676) receiving psychological treatment for depression. High rates of agreement in classification (mean agreement 94%) were found between four of the five methods (the Jacobson, NK, EN, and GLN methods). The HA method had the least agreement in classification with other methods (mean agreement 81%). Interestingly, the outcome measure used to calculate clinical significance impacted patient classification much more than the clinical significance method. This further highlights the need to agree on valid and reliable core outcome measures to be used within trials in a specific area. The authors concluded that these findings do not allow any recommendations to be made regarding which method to use.

4.6 The validity of different clinical significance classifications

The previous section focused on the convergences and divergences in classification rates between methods. Across studies, generally high rates of agreement in classification between methods were found. Because of this, and due to its wide-use and ease of computation, five of the six studies recommended the Jacobson method. However, a higher rate of agreement in classification does not demonstrate higher accuracy in classification. For instance, if five of the six methods demonstrate that around 60% of patients 'recovered', and the remaining method demonstrates that only 20% of patients 'recovered', this does not indicate that the latter method is of any less accuracy. It is important to consider whether methods are valid in their categorisation i.e. does a 'recovered' patient display characteristics expected from patients who have recovered. This is known as construct validity. Four empirical studies have investigated the construct validity of clinical significance methods - three focused solely on the Jacobson method, and one focused on both the Jacobson and HA methods. There findings will be discussed below.

Ankuta and Abeles (1993) categorised 74 outpatients with varied psychological disorders according to the Jacobson method and compared classifications with

patient's perceived satisfaction with therapy. Patients classed as 'recovered' reported higher levels of satisfaction than those not classed as 'recovered', providing important initial evidence of the construct validity of the Jacobson method.

Lunnen and Ogles (1998) extended the evaluation of Ankuta and Abeles (1993) by examining the construct validity of the Jacobson method from multiple perspectives. They categorised 52 outpatients with a range of psychological disorders according to the Jacobson method (RCI criterion only) and compared classifications from the three perspectives proposed by Strupp and Hadley (1977): the patient, the therapist, and a significant other. Perceived change, therapeutic alliance and satisfaction with services were evaluated from the three perspectives. From both the patient and therapist's perspective, patients classed as 'improved' showed higher levels of perceived change, therapeutic alliance, and satisfaction with therapy than those classed as 'unchanged' or 'deteriorated'. There were no significant differences between those who remained 'unchanged' or 'deteriorated' on any domain or from any perspectives. This demonstrates that the Jacobson method (RCI criterion) has good construct validity for defining patients as 'improved' but less construct validity for defining patients as 'deteriorated'.

Newnham, Harwood, and Page (2007) compared patient classifications according to the Jacobson method with self-rated QoL and therapist-rated functioning amongst 1,830 inpatients at a psychiatric hospital. Patients classed as 'recovered' and 'improved' had both higher self-rated QoL and therapist rated functioning than those classed as 'unchanged' or 'deteriorated', providing further evidence for the construct validity of the Jacobson method.

Ronk et al. (2016) evaluated the construct validity of the Jacobson and HA methods in classifying patients as 'recovered' amongst 119 patients recently discharged from a psychiatric hospital. 'Recovery' according to the Jacobson and HA methods were compared to three dimensions: consumer-based recovery, self-rated QoL, and readmission to a psychiatric unit. According to both clinical significance methods, patients classed as 'recovered' demonstrated higher levels of consumer-based recovery and self-rated QoL, and lower rates of hospital readmission than those not classed as 'recovered'. This suggests that both the Jacobson and HA methods have good construct validity for defining patients as 'recovered'. Since no meaningful

differences in capturing the construct of ‘recovery’ were found between methods, Ronk et al. (2016) supported the use of the Jacobson method due to its wide use and ease of computation.

4.7 Conclusion

While tests of statistical significance provide valuable group information, the practical relevance of this information is limited. An assessment of clinical significance is therefore needed to make meaningful inferences about the clinical impact of psychotherapy and provide a much clearer understanding of whether and how well a treatment ‘works’. The method of clinical significance proposed by Jacobson et al. (1984,1991) represents a meaningful and appropriate approach. While there are limitations of the Jacobson method, it has withstood vigorous debate and a superior method has yet to be established. It is also the most widely used method (Ogles et al., 2001). Therefore, the Jacobson method seems the most appropriate and valid method for measuring clinical significance in psychotherapy outcome research.

Chapter 5. Study 3: Do Manualised Psychological Treatments Alleviate Emotional Distress in Breast Cancer Patients? An Individual Patient Data Meta-Analysis

5.1 Introduction

Two Cochrane reviews (Jassim et al., 2015; Mustafa et al., 2013) and eight additional meta-analyses (Cobeanu & David, 2018; Duijts et al., 2011; Naaman et al., 2009; Tatrow & Montgomery, 2006; Xiao et al., 2017; Ye et al., 2018; Zhang et al., 2016; Zimmermann et al., 2007) of RCTs have concluded that efficacious psychological treatments for emotional distress in BCa exist. Consequently, healthcare policies and clinical practice guidelines internationally specify that psychological treatment should be available to BCa patients as part of their routine care throughout the disease trajectory (Dauchy et al., 2012; Holland et al., 2011; Howell et al., 2009; Li et al., 2016; National Breast Cancer Centre, 2003; National Comprehensive Cancer Network, 2003; National Institute for Clinical Excellence, 2004; Page & Adler, 2008; Reese et al., 2017; Tit et al., 2017).

However, study 2 challenged the conclusion that efficacious psychological treatments exist by exposing the poor quality of the trials. Two reasons why conclusions of previous meta-analyses are based on RCTs not relevant to clinical practice were identified. First, for a psychological treatment to be reproducible it must follow clear steps specified in a treatment manual (Westen, Novotny, & Thompson-Brenner, 2004). Therefore, the standards by which a treatment is considered "evidence-based" for inclusion in clinical practice in the United States require that a treatment be manualised (Chambless & Hollon, 1998). However, only 51% of RCTs in BCa used a manual and no meta-analysis excluded non-manualised treatments or examined whether manualisation influenced treatment efficacy. Second, clinical practice guidelines recommend psychological treatment only for BCa patients who are clinically distressed (National Comprehensive Cancer Network, 2003). However, only 15% of RCTs screened patients for distress, and no meta-analysis excluded non-distressed patients or provided separate analyses specifically for those with distress, limiting the generalisability of their conclusions to the clinical setting.

Moreover, previous meta-analyses are based solely on effect sizes. While effect sizes convey differences between conditions at the group level, they provide no information about individual variability in treatment response (Loerinc et al., 2015). This makes it difficult for researchers, clinicians, service providers and policy-makers to interpret the practical value of the treatments (Temple, Salmon, Tudur-Smith, Huntley, &

Fisher, 2018). To indicate the proportion of patients who benefit from treatment, and therefore its relevance to clinicians and services, an evaluation of clinical significance is needed. Unfortunately, in BCa, only 11% of RCTs and no meta-analyses have evaluated the clinical significance of treatments, making it difficult to interpret their practical value. When evaluating clinical significance, it is also important to assess 'deterioration' i.e. significant worsening in symptoms. In mental health settings, around 5-15% of patients deteriorate in psychotherapy RCTs (Rozental, Magnusson, Boettcher, Andersson, & Carlbring, 2017).

In addition, clinicians and service providers need two further kinds of information that meta-analyses have not yet provided. First, it is crucial to know the longer-term efficacy of psychological treatment (i.e. ≥6 months after completion of psychological treatment). Four previous meta-analyses evaluated longer-term effects (Cobeanu & David, 2018; Duijts et al., 2011; Jassim et al., 2015; Tatrow & Montgomery, 2006). However, three of these aggregated effects 0-12 months post-treatment (Duijts et al., 2011; Jassim et al., 2015; Tatrow & Montgomery, 2006), and one aggregated effects 3-12 months post-treatment (Cobeanu & David, 2018). Therefore, it is difficult to determine the durability of treatment gains. Second, whereas services need to know which types of psychological treatment are most efficacious, only three previous meta-analyses examined this, with each reaching a different conclusion; Jassim et al. (2015), Zimmermann et al. (2007), and Naaman et al. (2009) concluded that, relative to control conditions, CBT, psychoeducation and supportive therapy were the most efficacious treatments, respectively.

An individual patient data meta-analysis (IPD-MA) is therefore needed to overcome the limitations of previous meta-analyses. IPD-MAs are considered the 'gold standard' in meta-analysis techniques (Stewart & Parmar, 1993; Stewart & Tierney, 2002). Instead of using summary statistics from published RCTs, IPD-MA combines participant-level data from each relevant RCT into a common dataset. IPD-MA has three specific advantages over traditional meta-analysis (Stewart & Parmar, 1993). First, IPD-MA leads to more reliable analyses (Cooper & Patall, 2009; Stewart & Tierney, 2002; Tierney et al., 2015). Second, IPD-MA provides the opportunity for analyses which have not been reported in the included RCTs (Cooper & Patall, 2009). Crucially, this enables evaluation of treatment effects separately for distressed & non-distressed patients, yet to be done in BCa. Finally, IPD-MA is needed to evaluate the

clinical significance of treatments, a major omission from previous meta-analyses. As detailed in chapter 4, the clinical significance approach developed by Jacobson and colleagues (Jacobson et al., 1984; Jacobson & Truax, 1991) has become the most widely used method but has never been applied in this population.

The Jacobson method has two criteria. The first involves calculation of a cut-off point on a well-validated outcome measure to determine whether an individual's post-treatment score has a greater probability of being drawn from a functional or dysfunctional population. Second the 'reliable change index' (RCI) determines if the extent of change from pre- to post-treatment is statistically reliable. Applying the two criteria from the Jacobson method, patients in RCTs can be classed into one of four categories: i) 'recovered', if they make a statistically reliable change and move from a dysfunctional to a functional population; ii) 'improved', if they make a statistically reliable change but do not move from a dysfunctional to a functional population; iii) 'unchanged', if they do not make a statistically reliable change; and iv) 'deteriorated', if they make a statistically reliable change for the worse. Because most RCTs in BCa have been offered to patients irrespective of distress levels, many patients are likely to have been part of the functional population before treatment. Therefore, full assessment of clinical significance can only be applied to those who are distressed pre-treatment. It is, however, still possible to evaluate whether non-distressed patients satisfy Jacobson's second criterion and make a statistically reliable change. Therefore, these patients can be classed as 'improved', 'unchanged', or 'deteriorated'.

5.1.1 Aims of the present study

The aim of the present study was to conduct an IPD-MA to evaluate the efficacy of manualised psychological treatments for emotional distress in BCa. To detect whether different methods of analysis suggested different conclusions, efficacy was evaluated using both IPD effect size analysis and Jacobson's clinical significance analysis. These analyses were conducted for both the total sample (i.e. all patients irrespective of their pre-treatment distress levels) and the distressed sub-sample. For each analysis, efficacy was assessed at post-treatment and follow-up, and the influence of treatment type and methodological quality on outcome was investigated.

5.2 Method

This review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Liberati et al., 2009).

5.2.1 Eligibility criteria

Eligibility criteria follow the PICOS framework (Liberati et al., 2009).

Participants. Trials exclusively comprising adults aged 18 years or older with a histologically confirmed diagnosis of BCa.

Interventions. Psychological interventions were defined as manualised treatments (i.e. trials referring to the use of a manual to standardise treatment) using psychological techniques. Trials evaluating complementary alternative medicines (i.e. yoga, hypnosis, reiki, logotherapy, art therapy, dance therapy) or treatments involving no interaction between therapist and patient (i.e. based explicitly on written or visual material) were excluded.

Controls. Trials using either no treatment (usual care) or active (attention placebo) control groups. Trials comparing two or more specific psychological treatments without a control were excluded.

Outcomes. The primary outcome was emotional distress defined as anxiety, depression, or general distress. The use of a common outcome measure with well-established psychometric properties allows comparisons of clinically significant change between trials (Sheldrick, Kendall, & Heimberg, 2001). Therefore, included trials had to measure general distress using the Hospital Anxiety and Depression scale total (HADS-T; Zigmond & Snaith, 1983) or the Profile of Mood States Total Mood Disturbance (POMS-TMD; McNair, 1971), depression using the HADS depression subscale (HADS-D; Zigmond & Snaith, 1983) or the Centre for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977), and anxiety using the HADS anxiety subscale (HADS-A; Zigmond & Snaith, 1983). These measures are the most widely used (Temple et al., 2018) and well-validated (Hann, Winter, & Jacobsen, 1999; Johnston, Pollard, & Hennessey, 2000; Nyenhuis, Yamamoto, Luchetta, Terrien, & Parmentier, 1999) outcome measures in BCa RCTs.

Studies. Only RCTs published in English in a peer-reviewed journal.

5.2.2 Search strategy & study selection

PubMed, PsycINFO, Web of Science, Scopus, PsycARTICLE, and AMED were searched from their inception until 20th June 2018 using the same search strategy developed for study 2. The final search strategy used for each database is available in Appendix 2. Reference lists of both eligible RCTs and previous meta-analyses in BCa were hand-searched for additional studies.

Initially, titles and abstracts of all papers were screened by one reviewer (JT) to remove duplicates and clearly irrelevant studies. To check for consistency in selection, a random 50% of abstracts and titles were independently assessed by a second reviewer (AB). At this stage, agreement between reviewers was high (81%). Most disagreements were due to one reviewer adopting an overly liberal approach. Next, full text articles of potentially relevant papers were independently assessed for inclusion by both reviewers (JT & AB). At this stage, agreement between reviewers was also high (92%). Following assessment of inter-rater agreement, discrepancies were resolved through discussion with a third reviewer (PF).

5.2.3 Data extraction

The corresponding authors of eligible trials were contacted by email and raw data on relevant outcome measures at pre-treatment, post-treatment and follow-up, plus treatment condition was requested. If corresponding authors did not respond after two weeks, a reminder email was sent. If there was still no response after an additional two weeks, all authors named on the paper of the eligible trial were contacted by email. If none of the authors responded within a month, a reminder email was sent to all authors. If there was still no response after an additional month, it was assumed that the IPD was unobtainable.

Two reviewers (JT & CH) extracted data from published reports of eligible trials using a standardised data extraction protocol (see Appendix 3). Data extracted was year of publication; country of origin; number of patients randomly assigned to each condition; mean patient age in each condition; distribution of trajectory stage; distribution of tumour stage; choice of outcome measures; type of treatment and control conditions; mode of delivery of treatment and control conditions; and duration and number of sessions of treatment and control conditions.

5.2.4 Coding of treatment type

Treatments were coded into five broad categories*: 'cognitive-behavioural-based treatments'* (CBT; treatments targeting specific thoughts or behaviours using cognitive behavioural techniques); *'mindfulness-based treatments'* (MBT; treatments focusing on meditation, visualisation, and present-moment awareness); *'psycho-education'* (treatments primarily providing psychological education); *'support'* (treatments emphasising a supportive environment by providing emotional or social support); or *'other'* (treatments that either did not fit a defined category or combined different approaches without emphasising any one). The categorisation of treatments was discussed by three reviewers (JT, PF & CH) until consensus was reached. These categories matched the categories used in study 2.

5.2.5 Methodological quality

Methodological quality was assessed by two reviewers (JT & CH) using a modified version (Temple et al., 2018) of the POMRF (Öst, 2008). This consists of 19 items each scored 0 (poor), 1 (fair), or 2 (good), producing a total score ranging from 0 to 38, with higher scores indicating greater quality (see study 2 for more details on the POMRF). Interrater reliability between reviewers, assessed using the ICC, was 0.83 (95% CI, 0.51–0.94) indicating good inter-rater reliability. Following calculation of the ICC, discrepancies were resolved by discussion with a third reviewer (PF).

5.2.6 Identifying the distressed sub-sample

Patients were included in the distressed sub-sample who scored above established cut-offs for clinical levels of emotional distress pre-treatment: HADS-A (≥8; Hinz & Brähler, 2011); CES-D (≥16; Radloff, 1977); HADS-D (≥8; Hinz & Brähler, 2011); HADS-T (≥13; Hinz & Brähler, 2011); or POMS-TMD (≥37; Cella et al., 1989; Classen et al., 2008).

5.2.7 General analysis strategy

Each outcome variable (anxiety, depression, and general distress) was examined separately, and included outcome data measured using different relevant measurement tools (i.e. HADS-D & CES-D for depression; and HADS-T & POMS-TMD for general distress) in the same analysis. Outcomes were examined at two time points:

post-treatment, defined as the earliest assessment point ≤8 weeks after treatment ended; and follow-up, defined as the earliest assessment point ≥6 months after treatment ended (the original plan was to evaluate treatment effects ≥12 months after treatment ended but only one trial provided such data). Both the IPD effect size analysis and Jacobson's clinical significance analysis comprised three main elements:

Effect size analysis

1. Calculation of standardised mean difference (SMD) effect sizes in the total sample.
2. Calculation of SMD effect sizes specifically in the distressed sub-sample.
3. Evaluation of whether effect sizes varied with treatment type or methodological quality for the total sample and the distressed sub-sample.

Clinical significance analysis

1. Calculation of statistically reliable change in the total sample, using Jacobson's second criterion (RCI).
2. Calculation of statistically reliable change and recovery specifically amongst the distressed subsample, using Jacobson's first and second criteria.
3. Evaluation of whether the likelihood of achieving statistically reliable change or recovery varied with treatment type or methodological quality for the total sample and distressed sub-sample, respectively.

5.2.8 Statistical analysis

5.2.8.1 Preliminary analysis

Not all eligible RCTs provided IPD. Therefore, pooled effect sizes of trials providing IPD were compared with those not providing IPD at post-treatment to assess whether effects were consistent. For the RCTs not providing IPD, effect sizes and 95% CIs were calculated using the data available from published reports. Differences were tested using Cochrane's Q-test. Differences were also explored between trials providing IPD and those not providing IPD in sample characteristics, psychological treatment characteristics and methodological quality using chi-square tests and *t*-tests. Publication bias was assessed at post-treatment by inspecting funnel plots and using Egger's test (Egger, Smith, Schneider, & Minder, 1997).

5.2.8.2 Effect size analysis

Total sample. SMD effect sizes with 95% CIs were calculated using IPD by dividing the difference in mean value between treated and control patients by the pooled SD. SMDs were adjusted for small-sample bias using Hedges' g (Hedges, 1989) and were pooled across trials using the inverse variance random effects model (DerSimonian & Laird, 1986). All SMDs were scaled so that positive values represented effects in favour of treated patients. Between-group SMD effect sizes of 0.2, 0.5 and 0.8 were considered small, medium and large, respectively (Cohen, 1988). In cases where trials had multiple treatment or control groups, each comparison was evaluated separately. However, as multiple comparisons from the same trial are not mutually independent, the number of patients in the relevant treatment or control group was divided equally between each comparison. Heterogeneity between studies was tested using Cochrane Q-test and the proportion of total variation that was due to heterogeneity expressed as the I^2 statistic, with values greater than 50% indicating at least moderate heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003). Effect size analyses were conducted on a completer rather than an ITT basis because IPD was only received for treatment completers.

Distressed subsample. The SMD effect size analysis techniques were identical to those used for the total sample.

Influence of treatment type and methodological quality on treatment outcomes. These influences could only be explored at post-treatment because too few data were available at follow-up. Sub-group analyses were conducted using Cochrane's Q test to explore whether treatment type influenced effect sizes, and weighted regression analyses to explore whether methodological quality influenced effect sizes.

5.2.8.3 Clinical significance analysis

Total sample. Using the RCI formula presented on page 77 and the data in Table 5.1, the RCI (Jacobson's second criterion) was calculated separately for each outcome measure to determine the change in score that would be statistically reliable at $p<0.05$ (Table 5.1).

The proportions of patients classed as 'improved', 'unchanged', or 'deteriorated' following treatment, and 95% CIs, were calculated for each treatment and control

group. Proportions were pooled across trials using a random effects proportion meta-analysis (DerSimonian & Laird, 1986). The Stuart-Ord (inverse double arcsine square root) method (Stuart & Ord, 1994) was used to stabilise variance among studies.

Risk differences (RDs) with 95% CIs were then calculated to estimate the difference in improvement rates between treated and control patients. RDs were calculated by subtracting the proportion of patients who 'improved' in the control group from that in the treatment group. RDs were pooled across trials using the Mantel-Haenszel random effects model (Mantel & Haenszel, 1959). All RDs were scaled so that positive values represented effects in favour of treated patients. Tests of heterogeneity were conducted and trials with multiple comparisons were evaluated using the same procedures as in the effect size analysis. All RD analyses were conducted on a completer basis as explained above.

Distressed sub-sample. In addition to the RCI, a cut-off point to determine whether a patient was more likely to be drawn from a functional or dysfunctional population was calculated separately for each outcome measure. As explained in chapter 4, three methods exist to determine this cut-off (Jacobson et al., 1984): (a) patients' post-treatment score falls outside the range of dysfunctionality, defined as falling two or more SDs beyond the mean of the dysfunctional population, in the direction of functionality; (b) patients' post-treatment score is within the range of functionality, defined as falling within two SDs of the mean of the functional population; and (c) patients' post-treatment score has a greater probability of being drawn from the functional than the dysfunctional population (Jacobson et al., 1984). When normative data exists for both functional and dysfunctional populations, cut-off point (c) is the method of choice. Unfortunately, for the HADS, CES-D, and POMS, appropriate normative data for functional populations was not available. Thus, for the present study, only cut-off point (a) could be used. Table 5.1 shows the cut-off score required on each outcome measure to be within the range of the functional population.

The proportions of patients classed as 'recovered', 'improved', 'unchanged', or 'deteriorated' following treatment, and 95% CIs were calculated for each treatment and control group. Proportions were pooled across trials with the same techniques used for the total sample. Next, RDs with 95% CIs were calculated to estimate the difference in recovery rates between treated and control patients. RDs were calculated

by subtracting the incidence of 'recovery' in the control group from that in the treatment group. The model and procedures were identical to those used for the total sample analysis.

Influence of treatment type and methodological quality on treatment outcomes. The same techniques as in the effect size analysis were used to explore whether RDs were influenced by treatment type or methodological quality.

5.2.9 Statistical software

Effect size and RD analyses were conducted using comprehensive meta-analysis, version 3.3.07; while proportion meta-analyses were conducted using StatsDirect, version 3.0.171.

Table 5.1: Data used to determine the Reliable Change Index (RCI) and cut-off point on each outcome measure

Outcome measure	n_1	M_1	SD_1	S_{diff}	R_{xx}	RCI	Cut-off point
CES-D	351[a]	27.95	10.98	5.15	0.89	10	6
HADS-D	522[b]	10.52	2.59	1.68	0.79	3	5
POMS	397	66.83	24.75	11.07	0.9	22	17
HADS-T	817[b]	19.86	5.27	2.79	0.86	5	9
HADS-A	854[b]	11.45	2.9	1.74	0.82	3	6

Note. n_1= number of individuals scoring above established cut-offs for clinical levels of emotional distress pre-treatment; M_1= pre-treatment mean; SD_1 = pre-treatment standard deviation; Sdiff = pre-treatment standard error of difference; R_{xx} = internal consistency; RCI = reliable change index at $p<0.05$; [a] we included pre-treatment scores on the CES-D (n=39) from one additional trial (Badger et al., 2013) for which we received IPD but excluded from the IPD-MA because there was no control condition, and one additional treatment group from Stanton et al. (2005) for which we received IPD (n=43) but excluded from the IPD-MA because it involved no interaction between therapist and patient (i.e. it was based exclusively on visual material); [b] we included pre-treatment scores (n=76 for HADS-T & HADS-D; n=64 for HADS-A) from one additional trial (Rissanen, Nordin, Ahlgren, & Arving, 2015) for which we received IPD but excluded from the IPD-MA because there was no control condition

5.3 Results

5.3.1 Study selection

The database search retrieved 2,344 citations; 10 more were identified through hand searching. After removal of duplicates, 1,590 remained for screening based on title and abstract. Of these, 1,304 clearly did not meet inclusion criteria. Full text articles of the remaining 286 citations were retrieved and assessed. In total, 28 articles corresponding to 26 RCTs published between January 1980 and January 2018 were

eligible. Seventeen (n=2,996) of the 26 (n=5,049) eligible RCTs provided IPD and were included. Figure 5.1 shows the study selection. A complete list of references of the included RCTs can be found in Appendix 6.

5.3.2 Study and patient characteristics

Table 5.2 describes the included RCTs. In total, 20 treatment and 18 control groups were included in our analyses. Nine treatments were categorised as CBT, four as 'other', three as support, three as psychoeducation, and one as MBT. The mean duration of treatment was 14 hours (median 14; range 2-30) over nine sessions (median 9; range 2-18). Of the 18 control groups, six were categorised as TAU, six as active controls, three as WLCs, two as educational material, and one as assessment only. The duration and number of sessions of control groups was only reported in 12 trials. Of these, the mean duration was seven hours (median 6; range 0-30) over three sessions (median 1; range 0-10). The mean post-treatment assessment took place 1 week after treatment ended (median 0; range 0-8). Of the twelve trials reporting follow-up data (i.e. the earliest assessment point ≥6 months after treatment ended), the mean follow-up assessment took place 8 months after treatment ended (median 8; range 6-12). The mean total quality score on the POMRF was 16.5 out of 38, with median 16, and range 9-29 (i.e. 43% of the maximum possible score, median 42%, range 24-76%).

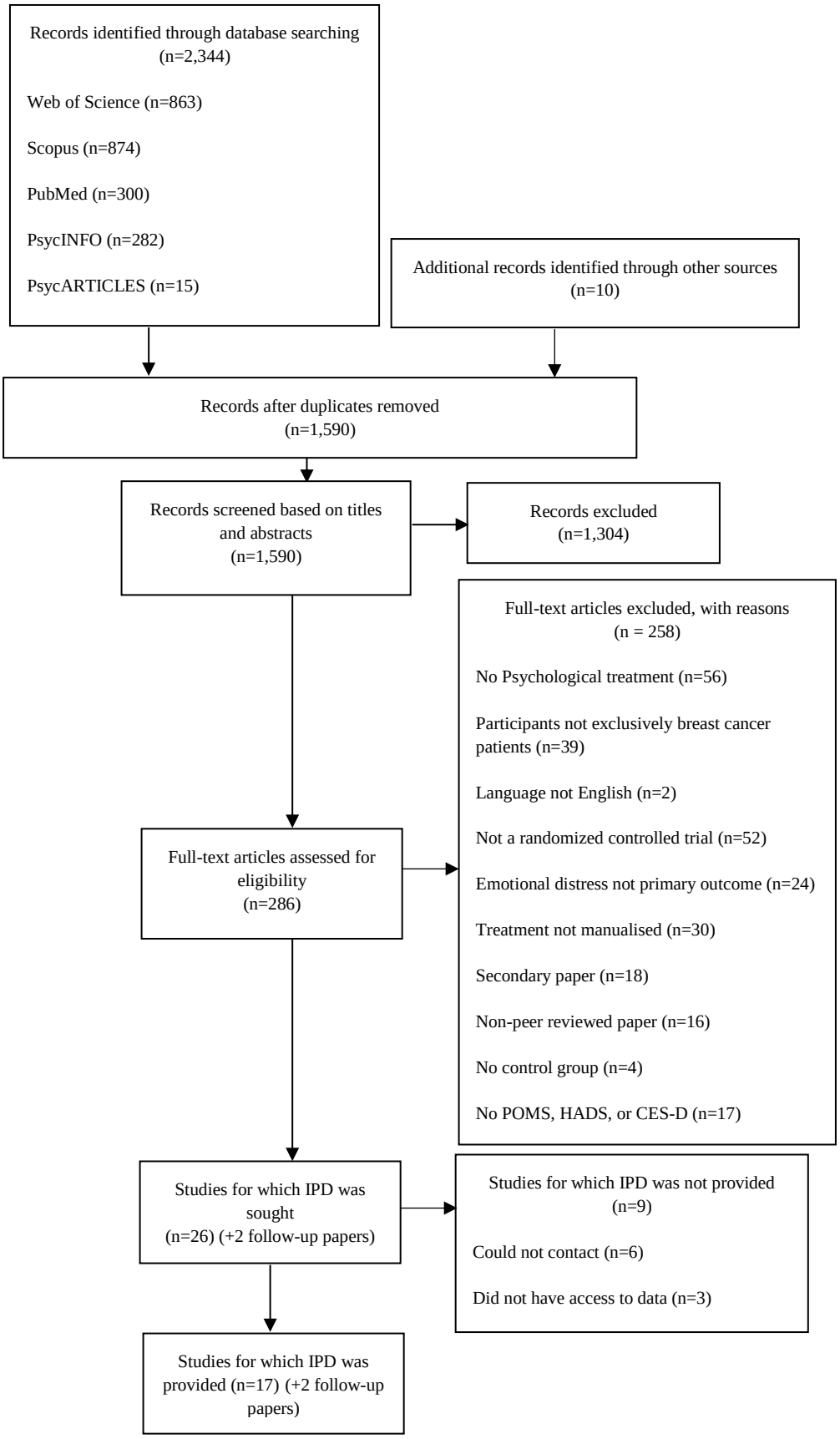


Figure 5.1: Flow chart showing trial identification and selection

Table 5.2: Characteristics of trials included in the IPD-MA

Author	Sample size	Treatment			Control			Outcome measures	Quality score
		Type of treatment	Mode of delivery	Total duration (hours)	Type of control	Mode of delivery	Total duration (hours)		
Andersen et al. (2004; 2007)	227	Other	Group	27	Assessment only	-	0	POMS	15
Antoni et al. (2001)	136	CBT	Group	20	Condensed version of active treatment[a]	Group	5-6	CES-D	14
Beutel et al. (2014)	156	Other	Individual	NR[b]	Offered referral to GP for psychological or pharmacological treatment[c]	-	NR[d]	HADS	20
Bredal et al. (2014)	367	Psychoeducation	Group	10	Nurse-led support[c]	Group	6	HADS	12
Carlson et al. (2013)	271	MBT	Group	21	Stress management seminar[a]	Group	6	POMS	19
		Supportive therapy	Group	18					
Classen et al. (2008)	353	Support + education material	Group	18	Education material	-	-	HADS, POMS[e]	16
Desautels et al. (2018)	62	CBT	Individual	8	Wait-list control	-	0	HADS	29
					Bright light therapy[a]	Individual	NR		

Author	Sample size	Treatment			Control			Outcome measures	Quality score
		Type of treatment	Mode of delivery	Total duration (hours)	Type of control	Mode of delivery	Total duration (hours)		
Dirksen et al. (2007)	81	CBT	Group	5.5[f]	Educational components of active treatment[a]	Group	5.5[f]	CES-D	16
Graves et al. (2003)	32	CBT	Group	12	TAU from medical community + educational material[c]	NR	NR	POMS	10
Groarke et al. (2012)	179	CBT	Group	15	Support from oncology nurses[c]	NR	NR	HADS	15
Ho et al. (2016)	157	Other	Group	16	Self-help support group[a]	Group	16	HADS	18
		Supportive therapy	Group	16					
Lechner et al. (2014)	114	CBT	Group	15	Educational information delivered by therapist[a]	Group	15	CES-D	19
Merckaert et al. (2016)	159	CBT	Group	30	Peer support[c]	Group	16	HADS	15
Sandgren et al. (2003; 2007)	237	Psychoeducation	Telephone	3	Nurse help line was available[c]	Telephone	NR	POMS	9
		Other	Telephone	3	-	-	-	-	
Savard et al. (2005)	57	CBT	Group	13.5	Wait-list control	-	0	HADS	19

Author	Sample size	Treatment			Control			Outcome measures	Quality score
		Type of treatment	Mode of delivery	Total duration (hours)	Type of control	Mode of delivery	Total duration (hours)		
Savard et al. (2006)	37	CBT	Individual	11	Wait-list control	-	0	HADS	20
Stanton et al. (2005)	371	Psychoeducation + education material	Individual	1.8	Education material	-	-	CES-D	15

Note. Hyphen indicates not applicable; TAU = treatment as usual CBT = cognitive behavioural therapy; MBT = mindfulness-based therapy; NR = not reported; [a] active control; [b] number of sessions varied (mean 18, range 0-31); [c] treatment as usual; [d] number of sessions varied (mean 2.4, range 0-24); [e] Classen et al. (2008) provided IPD for general distress on both the HADS-T and POMS-TMD. We chose to use IPD for general distress on the POMS-TMD as this was their primary outcome measure; [f] session 1 = 2 hours, sessions 2-4 = ≤1 hour, sessions 5-6 = 15 minutes

Of the 17 trials (n=2,966) for which IPD was obtained, 1,518 (treatment=775; control=743), 2,209 (treatment=1,119; control=1,090) and 2,267 (treatment=1,296; control=971) patients, respectively, provided outcome data for anxiety on the HADS-A (treatment=775; control=743), depression on the HADS-D (treatment=781; control=728) or CES-D (treatment=338; control=367), and general distress on the HADS-T (treatment=596; control=561) or POMS-TMD (treatment=700; control=410).

For patients providing outcome data for anxiety, the weighted mean dropout rate in the treatment group was 8% (range: 0-29%) at post-treatment and 6% (range: 0-20%) at follow-up. This compared to 9% (range: 0-26%) and 13% (range: 8-17%) in the control group, respectively. For patients providing outcome data for depression, the weighted mean dropout rate in the treatment group was 8% (range: 0-29%) at post-treatment and 9% (range: 0-27%) at follow-up. This compared to 9% (range: 0-26%) and 12% (range: 0-24%) in the control group, respectively. For patients providing outcome data for general distress, the weighted mean dropout rate in the treatment group was 12% (range: 0-53%) at post-treatment and 7% (range: 0-20) at follow-up. This compared to 10% (range: 0-65%) and 12% (range: 0-24%) in the control group, respectively.

Of the 2,966 patients for which IPD was obtained, 1,451 (49% of the total sample) scored above the established cut-offs for clinical levels of emotional distress pre-treatment and were therefore included in the distressed sub-sample. Of these patients, 790 (treatment=399; control=391), 749 (treatment=390; control=359) and 1,023 (treatment=590; control=433), respectively, scored above the established cut-offs for anxiety on the HADS-A (treatment=399; control=391), depression on the HADS-D (treatment=245; control=235) or CES-D (treatment=145; control=124), and general distress on the HADS-T (treatment=325; control=301) or POMS-TMD (treatment=265; control=132).

5.3.3 Preliminary analysis

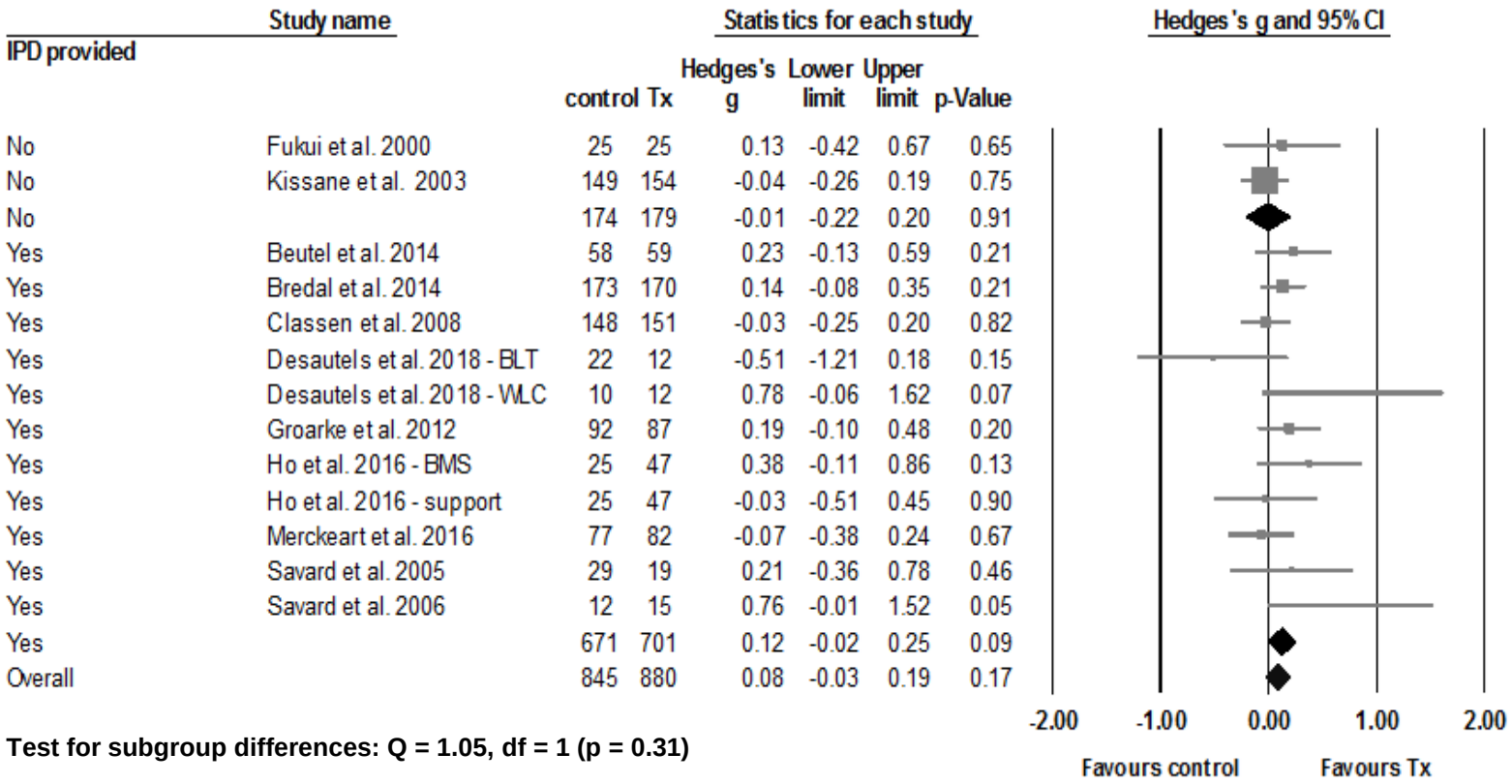
Trials that provided IPD (n=17) and those that did not provide IPD (n=9) did not differ on any outcome at post-treatment (see Figure 5.2, 5.3 & 5.4) nor did they differ on sample characteristics, psychological treatment characteristics or methodological quality (see Table 5.3). Thus, results of the IPD-MA are unlikely to be biased by

excluding trials not providing IPD. Inspection of funnel plots revealed little to no asymmetry (see Appendix 7); this was confirmed by Egger's regression test (all $p>0.05$).

5.3.4 Effect size analysis

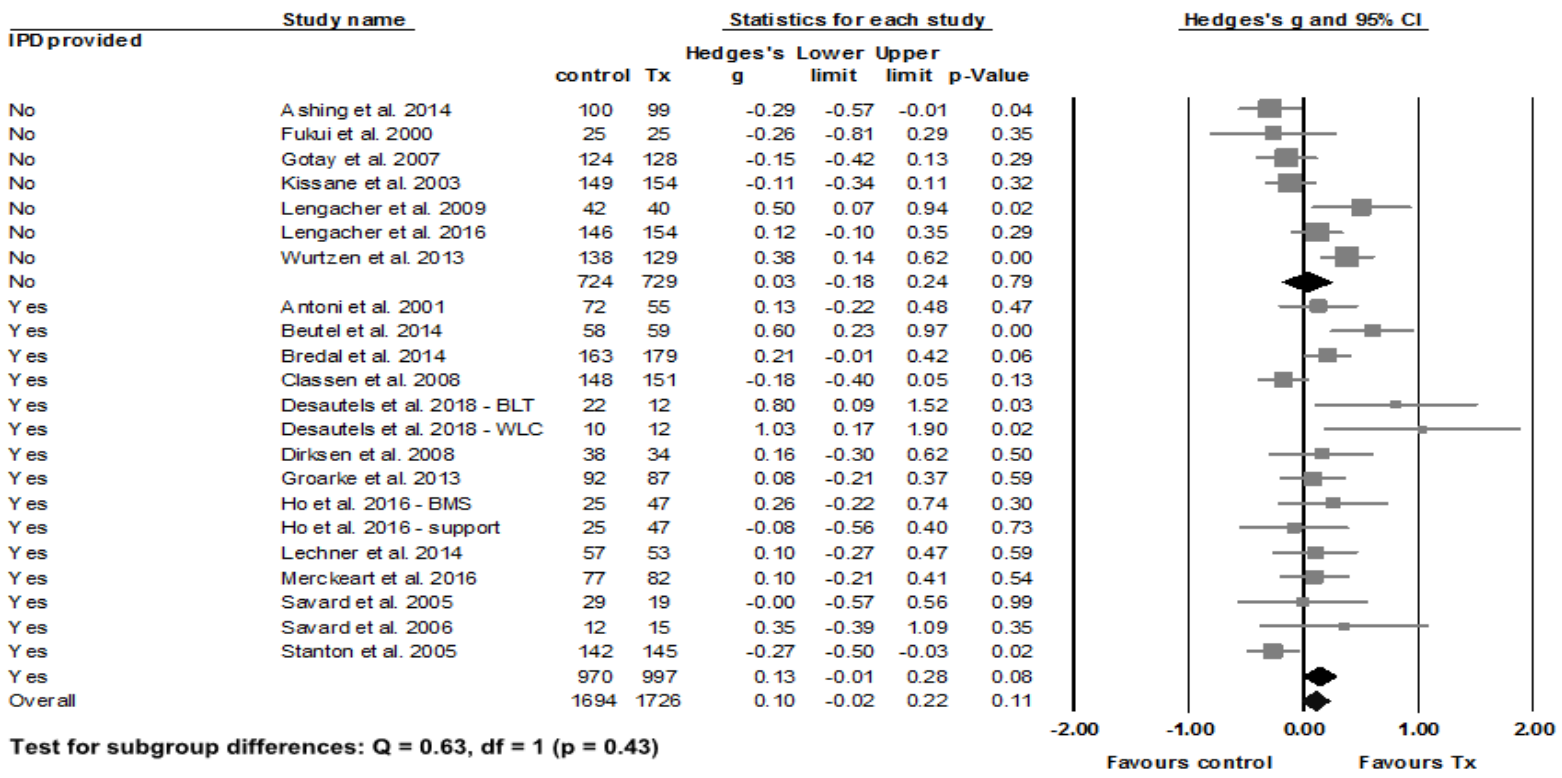
Total sample. At post-treatment and follow-up, the difference between treated and control patients, as measured by the pooled SMD effect size, was not statistically significant for anxiety, depression or general distress (see Table 5.4 & 5.6). Moderate heterogeneity was indicated for depression at post-treatment (i^2=58%) but not at follow-up. There was no evidence of heterogeneity for anxiety or general distress at post-treatment or at follow-up (see Appendix 8 for individual effect sizes for each trial).

Distressed subsample. At post-treatment, a small but significant effect size in favour of treated patients was found for anxiety (k=11, g=0.24, p=0.04), depression (k=15, g=0.33, p<0.01) and general distress (k=16, g=0.26, p<0.01; see Table 5.7). At follow-up, the difference between treated and control patients, as measured by the pooled SMD effect size, was no longer significantly different for anxiety, depression, or general distress (see Table 5.9). Moderate heterogeneity was indicated for anxiety at post-treatment (i^2=51%) but not at follow-up. There was no evidence of heterogeneity for depression or general distress at post-treatment or follow-up (see Appendix 9 for individual effect sizes for each trial).



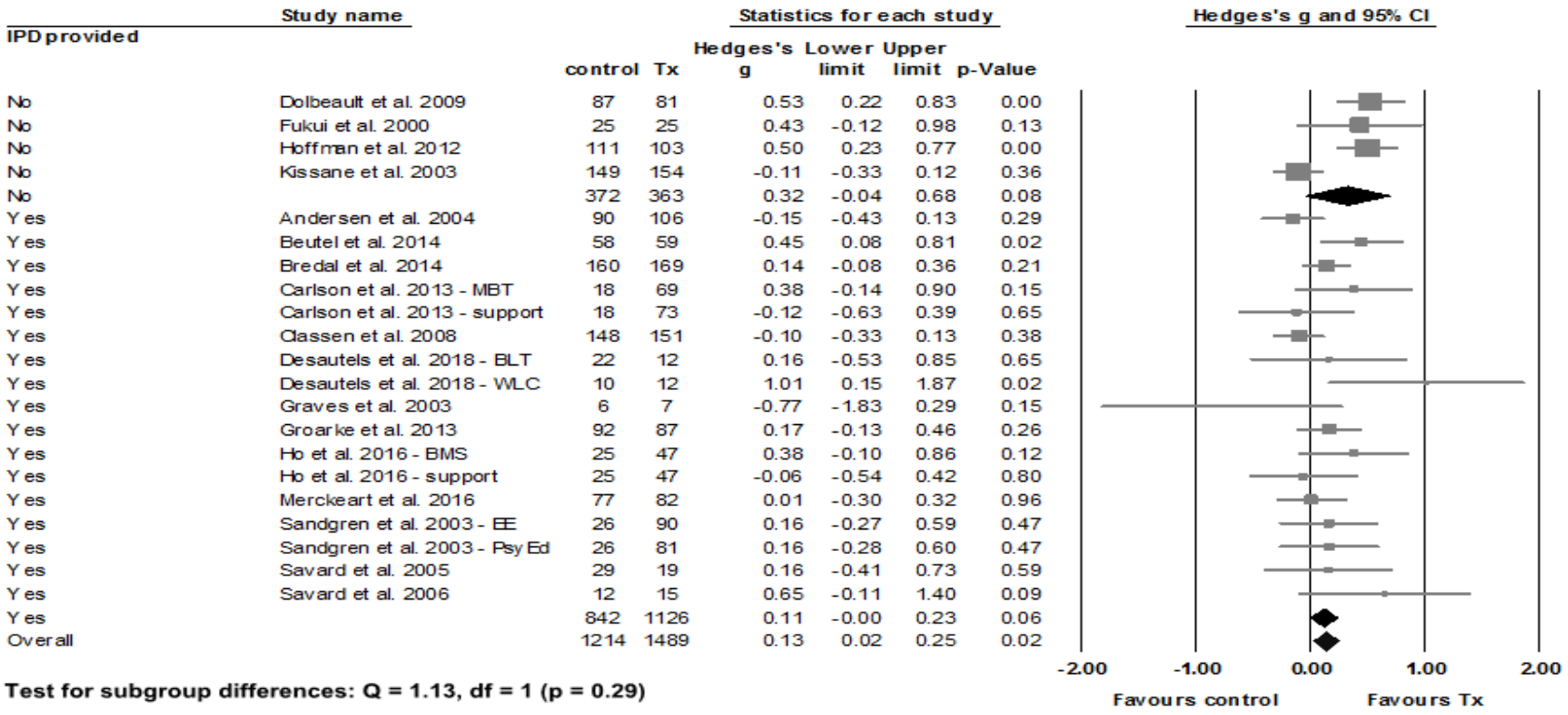
Note. BLT = bright light therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy

Figure 5.2: Comparison of effect sizes for trials providing IPD and those not providing IPD for anxiety at post-treatment



Note. BLT = bright light therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy

Figure 5.3: Comparison of effect sizes for trials providing IPD and those not providing IPD for depression at post-treatment



Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy; EE = emotional expression; PsyEd = psychoeducation

Figure 5.4: Comparison of effect sizes for trials providing IPD and those not providing IPD for general distress at post-treatment

Table 5.3: Comparison of sample characteristics, psychological treatment characteristics and methodological quality of trials providing IPD and those not providing IPD

Variable	IPD provided (n=17)	IPD not provided (n=9)	IPD provided vs IPD not provided	
			statistical test	p value
Sample characteristics				
Mean sample size	187	228	t(24)=-0.86	0.4
Mean age, years	53	53	t(24)=0.23	0.82
Psychological treatment characteristics				
Mean duration (hours)	14.4	15.5	t(26)=-0.5	0.62
Treatment type			$x^2(4)=8.27$	0.08
CBT	9	2		
Other	3	1		
Psychoeducation	3	0		
MBT	1	4		
Supportive therapy	4	2		
Treatment Format			$x^2(2)=0.34$	0.84
Individual	4	1		
Group	14	7		
Telephone	2	1		
Mean trial quality	16.5	14.3	t(24)=1.29	0.21

Note. CBT = cognitive behavioural therapy; MBT = mindfulness-based therapy; t = *t*-tests; x^2 = chi-square tests

Table 5.4: Treatment effects at post-treatment (≤8 weeks after treatment) in the total sample

	Effect sizes (g)								Risk differences for improvement							
Component	**k**	**n**	**g[a]**	**95% CI**	**p**	**i^2**	**Q (df)**	**p**	**k**	**n**	**RD[a]**	**95% CI**	**p**	**i^2**	**Q (df)**	**p**
Anxiety																
Overall effect	11	1,372	0.12	[-0.02,0.25]	0.09	0.25			11	1,370	0.05	[-0.04,0.14]	0.29	0.68		
Treatment type							3.23 (3)	0.36							1.77 (3)	0.62
CBT	6	469	0.16	[-0.14,0.45]	0.3	0.5			6	469	0.05	[-0.14,0.24]	0.61	0.8		
PsyEd	1	343	0.14	[-0.08,0.35]	0.21	0			1	343	0.03	[-0.07,0.13]	0.57	0		
Support	2	371	-0.03	[-0.23,0.18]	0.79	0			2	369	-0.01	[-0.09,0.07]	0.87	0		
Other	2	189	0.28	[-0.01,0.57]	0.06	0			2	189	0.11	[-0.05,0.27]	0.18	0.3		
Depression																
Overall effect	15	1,967	0.13	[-0.01,0.28]	0.08	0.58			15	1,954	0.06	[0.01,0.1]	0.02*	0.41		
Treatment type							11.96 (3)	0.01*							4.79 (3)	0.19
CBT	9	778	0.16	[0.02,0.31]	0.03*	0.03			9	777	0.08	[0.03,0.13]	<0.01*	0		
PsyEd	2	629	-0.03	[-0.49,0.44]	0.91	0.88			2	619	0	[-0.06,0.07]	0.92	0.36		
Support	2	371	-0.16	[-0.37,0.04]	0.12	0			2	371	0	[-0.07,0.08]	0.91	0		
Other	2	189	0.47	[0.14,0.79]	0.01*	0.19			2	187	0.16	[-0.15,0.48]	0.31	0.86		

	Effect sizes (g)								Risk differences for improvement							
Component	**k**	**n**	**g[a]**	**95% CI**	**p**	**i^2**	**Q (df)**	**p**	**k**	**n**	**RD[a]**	**95% CI**	**p**	**i^2**	**Q (df)**	**p**
							General distress									
Overall effect	17	1,968	0.11	[-0.00,0.23]	0.06	0.3			17	1,956	0.08	[0.03,0.13]	<0.01*	0.34		
Treatment type							6.02 (4)	0.2							7.69 (4)	0.1
CBT	7	482	0.18	[-0.07,0.43]	0.16	0.35			7	482	0.14	[0.07,0.21]	<0.01*	0		
MBT	1	87	0.38	[-0.14,0.9]	0.15	0			1	87	0.03	[-0.22,0.27]	0.82	0		
PsyEd	2	436	0.14	[-0.05,0.34]	0.15	0			2	427	0.01	[-0.08,0.1]	0.78	0		
Support	3	462	-0.1	[-0.29,0.09]	0.32	0			3	461	0.02	[-0.06,0.1]	0.62	0		
Other	4	501	0.18	[-0.12,0.49]	0.24	0.62			4	499	0.1	[-0.05,0.25]	0.19	0.69		

Note. k = number of treatment/control group comparisons; n = number of patients; PsyEd = psychoeducation; GP = general practitioner; g = hedges' g; RD = risk difference; CI = confidence interval; I^2 = measure of heterogeneity; Q = Cochran Q statistic; DF = degrees of freedom; CBT = cognitive behavioural therapy; MBT= mindfulness-based therapy; support = supportive therapy; [a] effect sizes and risk differences were scaled so that positive values represented effects in favour of treatment; * = p<0.05

Table 5.5: The influence of methodological quality on treatment effects at post-treatment (≤8 weeks after treatment) in the total sample

	Effect sizes (g)				Risk differences for improvement			
Moderator variable	**k**	**β**	**95% CI**	**p**	**k**	**β**	**95% CI**	**p**
			Anxiety					
Methodological quality	11	0.01	[-0.03,0.04]	0.79	11	-0.02	[-0.04,0.01]	0.08
			Depression					
Methodological quality	15	0.05	[0.01,0.09]	0.01*	15	0.01	[-0.01,0.04]	0.17
			General distress					
Methodological quality	17	0.02	[-0.01,0.05]	0.1	17	0.01	[-0.01,0.02]	0.06

Note. k = number of treatment/control group comparisons; g = hedges' g; RD = risk difference; CI = confidence interval; β = beta; * = p<0.05

Table 5.6: Treatment effects at follow-up (≥6 months after treatment) in the total sample

		Effect sizes (g)					Risk differences for improvement					
Outcome	**k**	**n**	**g[a]**	**95% CI**	**p**	**i^2**	**k**	**n**	**RD[a]**	**95% CI**	**p**	**i^2**
Anxiety	6	1,007	0.03	[-0.1,0.16]	0.64	0.02	6	1,005	0.02	[-0.04,0.08]	0.51	0.04
Depression	9	1,483	0.02	[-0.12,0.15]	0.79	0.35	9	1,470	0.01	[-0.04,0.05]	0.82	0.23
General distress	8	1,392	0.02	[-0.09, 0.13]	0.67	0.3	9	1,381	-0.01	[-0.05, 0.04]	0.77	0

Note. k = number of treatment/control group comparisons; n = number of patients; g = hedges' g; RD = risk difference; CI = confidence interval; I^2 = measure of heterogeneity; [a] effect sizes and risk differences were scaled so that positive values represented effects in favour of treatment

Table 5.7: Treatment effects at post-treatment (≤8 weeks after treatment) in the distressed sub-sample

	Effect sizes (g)								Risk differences for recovery							
Component	**k**	**n**	**g[a]**	**95% CI**	**p**	**i^2**	**Q (df)**	**p**	**k**	**n**	**RD[a]**	**95% CI**	**p**	**i^2**	**Q (df)**	**p**
Anxiety																
Overall effect	11	704	0.24	[0.00,0.47]	0.04*	0.51			11	704	0.13	[0.03,0.24]	0.01*	0.56		
Treatment type							3.71 (3)	0.3							2.51 (3)	0.47
CBT	6	231	0.44	[-0.06,0.94]	0.08	0.57			6	231	0.19	[-0.01,0.4]	0.07	0.5	0.71	
PsyEd	1	196	0.11	[-0.17,0.39]	0.45	0			1	196	0.04	[-0.09,0.17]	0.55	0	0	
Support	2	146	-0.05	[-0.37,0.28]	0.77	0			2	146	0.06	[-0.08,0.19]	0.3	0	0	
Other	2	131	0.32	[-0.0.,0.66]	0.07	0			2	131	0.15	[0.02,0.27]	0.02*	0.67	0	
Depression																
Overall effect	15	639	0.33	[0.13,0.53]	<0.01*	0.31			15	639	0.03	[-0.04,0.1]	0.39	0		
Treatment type							4.95 (3)	0.18							3.39 (3)	0.34
CBT	9	293	0.34	[0.11,0.5]	<0.01*	0			9	293	0.02	[-0.8,0.12]	0.67	0		
PsyEd	2	163	-0.01	[-0.32,0.3]	0.94	0			2	163	-0.01	[-0.14,0.12]	0.87	0		
Support	2	57	0.17	[-0.35,0.69]	0.54	0			2	57	-0.08	[-0.33,0.17]	0.53	0		
Other	2	126	1.12	[-0.22,2.46]	0.09	0.69			2	126	0.15	[-0.01,0.3]	0.06	0		

	Effect sizes (g)								Risk differences for recovery							
Component	**k**	**n**	**g[a]**	**95% CI**	**p**	**i^2**	**Q (df)**	**p**	**k**	**n**	**RD[a]**	**95% CI**	**p**	**i^2**	**Q (df)**	**p**
							General distress									
Overall effect	16	865	0.26	[0.08,0.44]	<0.01*	0.1			16	863	0.11	[0.06,0.17]	<0.01*	0.04		
Treatment type							8.22 (4)	0.08							6.05 (4)	0.2
CBT	6	215	0.56	[0.22,0.93]	<0.01*	0.04			6	215	0.21	[0.11,0.32]	<0.01*	0		
MBT	1	34	0.44	[-0.38,1.26]	0.29	0			1	34	0.05	[-0.36,0.47]	0.8	0		
PsyEd	2	194	0.12	[-0.16,0.41]	0.4	0			2	192	0.02	[-0.12,0.16]	0.76	0.07		
Support	3	163	-0.11	[-0.43,0.21]	0.5	0			3	163	0.04	[-0.13,0.22]	0.64	0.36		
Other	4	259	0.28	[-0.03,0.6]	0.08	0.09			4	259	0.11	[0.02,0.19]	0.01*	0		

Note. k = number of treatment/control group comparisons; n = number of patients; PsyEd = psychoeducation; g = hedges' g; RD = risk difference; CI = confidence interval; I^2 = measure of heterogeneity; Q = ratio of variation to within-study error; DF = degrees of freedom; CBT = cognitive behavioural therapy; MBT= mindfulness-based therapy; support = supportive therapy; [a] effect sizes and risk differences were scaled so that positive values represented effects in favour of treatment; * = $p<0.05$

Table 5.8: The influence of methodological quality on treatment outcome at post-treatment (≤8 weeks after treatment) in the distressed sub-sample

	Effect sizes (g)				Risk differences for recovery			
Moderator variable	**k**	**β**	**95% CI**	**p**	**k**	**β**	**95% CI**	**p**
			Anxiety					
Methodological quality	11	0.01	[-0.04,0.06]	0.65	11	-0.01	[-0.03,0.02]	0.5
			Depression					
Methodological quality	15	0.04	[-0.01,0.08]	0.08	15	0.01	[-0.01,0.03]	0.36
			General distress					
Methodological quality	16	0.03	[-0.01,0.07]	0.05	16	0.01	[-0.01,0.02]	0.16

Note. k = number of treatment/control group comparisons; g = hedges' g; RD = risk difference; CI = confidence interval; β = beta value

Table 5.9: Treatment effects at follow-up (≥6 months after treatment) in the distressed sub-sample

	Effect sizes (g)						Risk differences for recovery					
Outcome	**k**	**n**	**g[a]**	**95% CI**	**p**	**i^2**	**k**	**n**	**RD[a]**	**95% CI**	**p**	**i^2**
Anxiety	6	457	-0.07	[-0.25,0.12]	0.48	0	6	457	-0.04	[-0.12,0.05]	0.41	0
Depression	9	423	0.04	[-0.23,0.32]	0.76	0.45	9	423	-0.02	[-0.13,0.09]	0.71	0.23
General distress	9	548	0.13	[-0.04, 0.29]	0.15	0	9	548	0.03	[-0.03, 0.09]	0.36	0

Note. k = number of treatment/control group comparisons; n = number of patients; g = hedges' g; RD = risk difference; CI = confidence interval; I^2 = measure of heterogeneity; [a] effect sizes and risk differences were scaled so that positive values represented effects in favour of treatment

Influence of treatment type and methodological quality on treatment outcome. In the total sample, neither treatment type nor methodological quality was significantly related to variation in treatment effects for anxiety or general distress. Treatment type was significantly related to variation in treatment effects for depression (k=15, Q=11.96, df=3, p=0.01): significant effects in favour of treated patients were found for CBT (k=9, g=0.16, p=0.03) and 'other' treatments (k=2, g=0.47, p=0.01) but not

psychoeducation (k=2, g=-0.03, p=0.91) or support (k=2, g=-0.16, p=0.12). Methodological quality was also related to variation in treatment effects for depression (k=15, β =0.05, p=0.01), with higher quality studies having larger effects (see Table 5.4 & 5.5). Amongst the distressed subsample, neither treatment type nor methodological quality was related to variation in effect sizes for anxiety, depression, or general distress (see Table 5.7 & 5.8).

Summary. In the total sample, no treatment effects were found for anxiety, depression or general distress at post-treatment or follow-up. In the distressed sub-sample, treatment effects were found for anxiety, depression and general distress at post-treatment but not at follow-up. The only influence of treatment type and methodological quality on outcome was for depression in the total sample.

5.3.5 Clinical significance analysis

Total sample. Table 5.10 shows the pooled proportion of treated and control patients allocated to each reliable change category for each outcome variable at post-treatment and follow-up (see Appendix 10 for proportions for each trial). At post-treatment, across outcomes, few treated and control patients 'deteriorated' (treatment=9-11%; control=12-15%), a large minority 'improved' (treatment=26-34%; control=19-27%), and most remained 'unchanged' (treatment=55-64%; control=58-69%). At follow-up, the proportion of treated and control patients classed as 'deteriorated' was slightly higher than the proportions at post-treatment (treatment=9-21%, control=11-18%). The proportion of treated patients classed as 'unchanged' (52-68%) and 'improved' (24-34%) were comparable to post-treatment. However, compared to post-treatment, fewer control patients were classed as 'unchanged' (52-65%) and more as 'improved' (23-35%).

At post-treatment, a significant RD in favour of treated patients was found for depression (k=15, RD=0.06, p=0.02) and general distress (k=17, RD=0.08, p<0.01). This indicates that, on average, 6% and 8% more treated patients compared to control patients 'improved' for depression and general distress, respectively. The RD comparing improvement between treated and control patients was not statistically significant for anxiety (see Table 5.4). At follow-up, RDs comparing improvement between treated and control patients were not significantly different for anxiety, depression or general distress (see Table 5.6). Moderate heterogeneity was indicated

for anxiety at post-treatment (i^2=68%) but not at follow-up. There was no evidence of heterogeneity for depression or general distress at post-treatment or follow-up (see Appendix 11 for individual RDs for improvement for each trial).

Distressed subsample. Table 5.11 shows the pooled proportion of treated and control patients allocated to each Jacobson outcome category at post-treatment and follow-up (see Appendix 12 for proportions for each trial). At post-treatment, across outcomes, a small minority of treated and control patients 'deteriorated' (treatment=6-8%; control=7-12%), a larger minority 'improved' (treatment=17-26%; control=16-24%), around one quarter to one third 'recovered' (treatment=28-32%; control=17-27%), and a large proportion remained 'unchanged' (treatment=37-45%; control=50-51%). At follow-up, the proportions of treated and control patients who 'deteriorated' (treatment=5-11%; control=8-12%) were similar to those at post-treatment. The proportion of treated patients who remained 'unchanged' (41-47%), 'improved' (15-24%), and 'recovered' (24-33%) were also similar to those at post-treatment. However, compared to post-treatment, fewer control patients remained 'unchanged', (36-43%), and more 'improved' (15-28%) and 'recovered' (24-34%).

At post-treatment, a significant RD in favour of treated patients was found for general distress (k=16, RD=0.11, p<0.01) and anxiety (k=11, RD=0.13, p=0.01). This indicates that, on average, 11% and 13% more treated patients compared to control patients recovered from general distress and anxiety, respectively. The RD comparing recovery between treated and control patients was not significantly different for depression (see Table 5.7). At follow-up, RDs comparing recovery between treated and control patients were not significantly different for anxiety, depression, or general distress (see Table 5.9). Moderate heterogeneity was indicated for anxiety at post-treatment (i2=56%) but not at follow-up. There was no evidence of heterogeneity for depression or general distress at post-treatment or follow-up (see Appendix 13 for individual RDs for recovery for each trial).

Influence of treatment type and methodological quality. Neither treatment type nor methodological quality was significantly related to variation in RDs for improvement in the total sample (see Table 5.4 & 5.5) or recovery amongst the distressed sub-sample (see Table 5.7 & 5.8).

Summary. In the total sample, by contrast with the effect size analysis, significant RDs for improvement in favour of treated patients were found for depression and general distress at post-treatment. However, consistent with the effect size analysis, RDs for improvement were not statistically significant at follow-up. Similarly, in the distressed sub-sample, significant RDs for recovery in favour of treated patients for anxiety and general distress at post-treatment did not persist to follow-up. Neither treatment type nor methodological quality influenced RDs.

Table 5.10: Classification of patients at post-treatment (≤8 weeks after treatment) and follow-up (≥6 months after treatment) according to Jacobson's second criterion (RCI) in the total sample[a]

	Post-treatment (%)				Follow-up (%)			
Condition	n	Deteriorated	Unchanged	Improved	n	Deteriorated	Unchanged	Improved
Anxiety								
Treatment	700	11	58	31	524	14	52	34
95% CI		(7-15)	(50-66)	(23-39)		(10-18)	(40-64)	(20-50)
Control	672	15	58	27	481	13	52	35
95% CI		(11-19)	(49-67)	(19-36)		(9-18)	(40-63)	(22-51)
Depression								
Treatment	992	9	64	26	757	9	68	24
95% CI		(6-13)	(57-72)	(18-35)		(7-11)	(58-77)	(16-33)
Control	964	12	69	19	713	11	65	23
95% CI		(9-14)	(60-76)	(14-25)		(9-14)	(54-76)	(14-34)
General distress								
Treatment	1122	11	55	34	787	21	52	27
95% CI		(8-14)	(48-62)	(27-42)		(8-35)	(40-62)	(15-39)
Control	837	12	62	26	596	18	52	30
95% CI		(9-15)	(56-69)	(19-33)		(7-29)	(42-60)	(17-42)

Note. n = number of patients; RCI = reliable change index at P<0.05; CI = confidence interval; [a] percentages have been calculated by pooling arm level data across trials to provide information about the actual percentage estimates. However, to interpret direct comparative results of improvement rates between treatment and control groups the reader should refer to the risk difference analyses presented in tables 5.3 & 5.5

Table 5.11: Classification of patients at post-treatment (≤8 weeks after treatment) and follow-up (≥6 months after treatment) according to Jacobson's first and second (RCI) criteria in the distressed sub-sample[a]

	Post-treatment (%)					Follow-up (%)				
Condition	n	Deteriorated	Unchanged	Improved	Recovered	n	Deteriorated	Unchanged	Improved	Recovered
Anxiety										
Treatment	353	8	44	17	32	218	11	41	18	31
95% CI		(4-12)	(36-51)	(13-22)	(25-39)		(6-17)	(25-57)	(13-25)	(18-47)
Control	353	12	51	18	19	211	9	43	15	34
95% CI		(8-14)	(37-62)	(11-24)	(13-27)		(4-14)	(30-56)	(10-20)	(21-49)
Depression										
Treatment	334	6	45	18	32	225	5	47	15	33
95% CI		(3-8)	(34-56)	(12-25)	(20-45)		(2-7)	(32-62)	(7-25)	(12-58)
Control	305	7	51	16	27	199	8	41	19	31
95% CI		(4-11)	(45-60)	(11-21)	(16-39)		(4-12)	(23-56)	(11-27)	(9-59)
General distress										
Treatment	497	8	37	26	28	310	8	44	24	24
95% CI		(6-12)	(30-45)	(19-33)	(24-33)		(3-12)	(32-55)	(17-19)	(11-39)
Control	369	10	50	24	17	235	12	36	28	24
95% CI		(7-13)	(42-58)	(17-33)	(11-24)		(5-20)	(24-48)	(22-33)	(9-40)

Note. n = number of patients; RCI = reliable change index at $p<0.05$; CI = confidence interval; [a] percentages have been calculated by pooling arm level data across trials to provide information about the actual percentage estimates. However, to interpret direct comparative results of recovery rates between treatment and control groups the reader should refer to the risk difference analyses presented in tables 5.6 & 5.8.

5.4 Discussion

An IPD-MA was conducted to evaluate the efficacy of manualised psychological treatments for emotional distress in BCa in two different ways: effect size analysis and clinical significance analysis. When all patients were included in the analysis, irrespective of their levels of distress before treatment, the two methods converged on a disappointing picture. Despite evidence from the clinical significance analysis of benefits for depression and general distress immediately post-treatment, no advantage from treatment remained relative to control conditions after a mean of 8 months follow-up. When the subset of patients who were clinically distressed before treatment were examined, findings remained disappointing. Despite evidence from both analyses of benefits for anxiety and general distress at post-treatment, and evidence from the effect size analysis of benefits also for depression at post-treatment, there were no benefits for treatment relative to control conditions for any form of emotional distress at follow-up.

While the two methods of analysis largely converged, findings from the clinical significance analysis added a realistic indication of the *practical* significance of *statistically* significant findings. Despite statistically significant evidence that treatment improved outcomes at post-treatment, the clinical significance analysis showed that benefits were very small. In the total sample, only 6-8% more treated than control patients improved on measures of depression and general distress. In the distressed sub-sample, only 11-13% more treated than control patients recovered on measures of anxiety and general distress. Overall, regardless of whether patients received treatment or control conditions, the proportion recovering was low; across outcomes at post-treatment, only 28-32% of treated patients recovered compared to 17-27% of controls At follow-up, the proportion recovering remained low: only 24-31% of treated patients and 24-34% of controls recovered.

Apart from isolated findings in the effect size analysis for depression in the total sample; that higher-quality trials had larger effects, and that CBT and 'other' treatments were the only treatment types to have significant effect sizes favouring treated over control patients, neither methodological quality nor treatment type

influenced treatment outcome. As these findings were not borne out in the clinical significance analysis, their robustness is questionable.

Findings from the effect size analysis in the total sample at post-treatment can be directly compared to those of previous meta-analyses. Findings presented here differ from those (Cobeanu & David, 2018; Duijts et al., 2011; Jassim et al., 2015; Mustafa et al., 2013; Naaman et al., 2009; Tatrow & Montgomery, 2006; Xiao et al., 2017; Ye et al., 2018; Zhang et al., 2016; Zimmermann et al., 2007). Previous meta-analyses found significant controlled effect sizes favouring treated patients relative to control patients, ranging from small to large (0.26-1.11). However, controlled effect sizes comparing treated and control patients in the analysis presented in this study were non-significant. This divergence might reflect the type of RCTs included. By contrast with previous meta-analyses, only manualised psychological treatments were included in this study. Moreover, the divergence of conclusions drawn from this IPD-MA and those of previous meta-analyses reflects additional ways in which methods differed. Specifically, by contrast with previous meta-analyses, in this IPD-MA the clinical significance of psychological treatments was evaluated, and their longer-term efficacy was evaluated. Separate analyses were also provided for patients with clinical levels of emotional distress - the group for whom clinical practice guidelines explicitly recommend psychological treatment (National Comprehensive Cancer Network, 2003). Based on these crucial methodological differences, the conclusion of previous meta-analyses that efficacious psychological treatments for emotional distress in BCa exist cannot be supported, as these conclusions are based on patients and findings not relevant to clinical practice.

This meta-analysis has limitations. First, not all eligible RCTs were included as authors of only 17 of the 26 eligible RCTs provided IPD, potentially compromising the generalisability of the results. However, no differences were found in effect sizes between the 17 included RCTs and the 9 not included, suggesting that the included RCTs were representative of published trials. Second, the included RCTs reported limited long-term follow-up data. Therefore, the longer-term benefits of treatment could only be considered around 8 months after treatment completion. Third, RCTs not published in English language were excluded. Therefore, relevant RCTs may have been omitted. Fourth, moderate heterogeneity was identified for anxiety at post-

treatment which could not be explained by methodological quality or treatment type. Thus, heterogeneity must be due to other characteristics which were not explored. For instance, included RCTs were diverse with regards to the point in the disease trajectory at which patients received psychological treatment (i.e. soon after diagnosis, during medical treatment, or in survivorship). When requesting IPD, this information was not asked for and it could not be extracted from published reports because most RCTs aggregated patients at different points in the disease trajectory. Finally, this IPD-MA may have overestimated the efficacy of psychological treatments because analyses were only based on treatment completers, discounting patients who may have been non-compliant to treatment because of treatment failure. Notably larger effects have been found for trials conducted on a completer basis compared to an ITT basis (Cuijpers et al., 2010). However, dropout rates in this IPD-MA were relatively small (~10%) and distributed equally across conditions.

Healthcare policies currently specify that psychological treatment should be available to BCa patients as part of their routine care (Holland et al., 2011; Page & Adler, 2008; Tit et al., 2017). However, in light of the findings in this IPD-MA of the minimal and temporary benefit of treatment, its clinical utility is questionable. If BCa patients were aware of the small benefit psychological treatment offers, they might choose not to commit to psychotherapy, especially considering the commitment of time and personal involvement it requires. It could be argued that there is little point in providing BCa patients with these psychological treatments when other options, such as support from oncology nurses, are just as beneficial.

One possible explanation for these poor outcomes is that treatments were inadequately implemented. Although only trials that used manualised treatments were included, most trials did not use certified trained therapists or monitor therapists' adherence or competence to treatment. These design elements are essential to ensure that treatment was implemented as designed (Temple et al., 2018).

An alternative explanation for these poor outcomes might be that current therapeutic approaches are unsuitable for BCa patients. Nine of the 20 treatments were based on CBT, of which the central premise is that unrealistic appraisals of events initiate and maintain emotional distress. CBT therefore commonly seeks to reduce distress by

testing the reality of unrealistic thoughts (Bennett-Levy et al., 2004). However, thoughts and concerns troubling BCa patients are often not unrealistic (e.g. 'my cancer may return' or 'my cancer is putting a financial strain on my family'). Thus, challenging such thoughts may be inappropriate in BCa. Evaluating alternative psychotherapeutic approaches that do not focus on challenging the content of negative thoughts but instead the processes which lead individuals to respond negatively to such thoughts may be more suitable in BCa. An additional advantage of focusing on psychological processes is that it offers the potential to address comorbid problems. This could benefit BCa patients, who often present with mixed symptoms of anxiety, depression and general distress.

In conclusion**,** approximately 70% of BCa patients remain distressed following completion of psychological treatment. While patients receiving them are slightly more likely to recover than those receiving a control condition in the immediate short-term, they are no more likely to recover after around 8 months. More efficacious psychological treatments are urgently needed for BCa patients with emotional distress.

Chapter 6. Alternative Psychotherapeutic Approaches for Understanding Emotional Distress in Breast Cancer - The Intolerance of Uncertainty Model and the Self-Regulatory Executive Function (S-REF) Model

6.1 Introduction

Study 3 showed that current psychological treatments do not alleviate emotional distress in most BCa patients. Most of the psychological treatments evaluated in study 3 were either non-theory based or based on the central premise of the cognitive model (Beck, 1967, 1976), that is, that individual's appraisal of a situation, rather than the situation itself, determines the behaviours and emotions that follow. In other words, an event leads to a thought (or appraisal), which leads to an emotional response. Thus, according to the cognitive model, negative appraisals of situations and events (e.g. negative thoughts about BCa and its consequences) initiate and maintain emotional distress. Negative thoughts about BCa and its consequences are associated with current (Cook et al., 2015a) and future distress in BCa patients (Cook et al., 2015b; Millar et al., 2005). However, most if not all BCa patients experience negative thoughts. Yet, not all BCa patients become distressed. Thus, psychological mechanisms not accounted for by the cognitive model may underlie emotional distress in BCa patients.

Two potential mechanisms underlying emotional distress in BCa are worry and rumination, both of which are common among BCa patients. Worry has been defined as *"a chain of thoughts and images, negatively affect-laden, and relatively uncontrollable"* (Borkovec, Robinson, Pruzinsky, & DePree, 1983, p. 10) and has typically been examined in the anxiety literature. Rumination has been defined as *"behaviors and thoughts that focus one's attention on one's depressive symptoms and on the implications of these symptoms"* (Nolen-Hoeksema, 1991, p. 569) and has typically been examined in the depression literature. The main distinction between worry and rumination has been their temporal focus (i.e. worry is more future-oriented and rumination is more past-oriented; Borkovec, Ray, & Stober, 1998; Papageorgiou & Wells, 2004). However, both types of repetitive thinking share common processes and are highly correlated (Fresco, Frankel, Mennin, Turk, & Heimberg, 2002; Muris, Roelofs, Rassin, Franken, & Mayer, 2005; Segerstrom, Tsao, Alden, & Craske, 2000). Worry and rumination are also similarly related to anxiety and depression (D'Hudson & Saling, 2010; Fresco et al., 2002; Goring & Papageorgiou, 2008; Nolen-Hoeksema, 2000; Rood, Roelofs, Bögels, & Alloy, 2010; Segerstrom et al., 2000). Consequently, worry and rumination have been conceptualised as part of a broader transdiagnostic

construct labelled ‘repetitive negative thinking’ (RNT; Harvey, Watkins, & Mansell, 2004). RNT is associated with emotional distress in numerous clinical and non-clinical populations (Mahoney, McEvoy, & Moulds, 2012; McEvoy, Mahoney, & Moulds, 2010; McEvoy et al., 2017; Spinhoven, Drost, van Hemert, & Penninx, 2015); and RNT in the form of worry and rumination is associated with emotional distress in BCa (Chan, Ho, Tedeschi, & Leung, 2011; Gibbons, Groarke, & Sweeney, 2016; Lo-Fo-Wong et al., 2016; Soo & Sherman, 2015). However, the cognitive model, which most treatments in BCa are based on, does not attempt to explain what causes and maintains RNT.

Two psychological models that could account for RNT and subsequent emotional distress in BCa are the intolerance of uncertainty (IU) model (Dugas et al., 1998) and the Self-Regulatory Executive Function (S-REF) model (Wells & Matthews, 1994, 1996).

6.2 Intolerance of uncertainty model

The IU model (Dugas et al., 1998) was developed to explain the persistence of worry in generalised anxiety disorder (GAD). More recently, the IU model has been conceptualised as a transdiagnostic model which explains the development and maintenance of RNT and emotional distress across a range of mental and physical health populations (Mahoney & McEvoy, 2012; McEvoy & Mahoney, 2012).

The IU model has four components: IU; positive beliefs about worry; negative problem orientation; and cognitive avoidance. IU, the core component of the model, refers to a dispositional characteristic resulting from dysfunctional cognitive beliefs surrounding uncertain situations (Dugas, Buhr, & Ladouceur, 2004). Examples of these beliefs include “uncertainty makes me anxious”, “uncertainty stops me from functioning”, and “uncertainty makes life intolerable”. Individuals with high levels of IU perceive future uncertain events as unacceptable and intolerable, regardless of the probability of the event.

The IU model posits that IU directly leads to RNT by enhancing cognitive interpretational biases whereas the other three components (positive beliefs about worry; negative problem orientation; and cognitive avoidance) contribute to RNT via

indirect pathways. Positive beliefs about worry refer to beliefs that worrying will enable coping and prevent the occurrence of future unwanted events (e.g. "worrying will help me stop an event from occurring"; Dugas et al., 1998; Freeston, Rhéaume, Letarte, Dugas, & Ladouceur, 1994; Koerner & Dugas, 2006). Positive beliefs about worry can be maintained and strengthened through both positive reinforcement (e.g. worrying provides a solution to an uncertain situation) and negative reinforcement (e.g. worrying is followed by the absence of an uncertain situation occurring), leading to enhanced levels of RNT.

RNT and subsequent distress leads to negative problem orientation and cognitive avoidance (Behar, DiMarco, Hekler, Mohlman, & Staples, 2009). Negative problem orientation refers to a lack of problem-solving confidence. Negative problem orientation leads to limited attempts to deal with difficult situations, which perpetuates false beliefs about problem-solving abilities and exacerbates RNT. Cognitive avoidance, drawn from Borkovecs's avoidance theory (Borkovec, 1994; Borkovec, Alcaine, & Behar, 2004) refers to a variety of automatic and cognitive strategies to avoid or supress distressing thoughts and images. Examples of these strategies in BCa include supressing thoughts about BCa recurrence, replacing negative images of BCa with positive or neutral images, and avoiding BCa follow-up appointments. These strategies ultimately backfire and maintain and exacerbate RNT and subsequent distress by enhancing the strength and frequency of distressing thoughts and images (Borkovec & Roemer, 1995; Dugas et al., 1998).

6.2.1 Empirical support for the intolerance of uncertainty model

IU significantly predicts emotional distress and RNT in the form of worry and rumination in a range of clinical populations (Carleton, 2012; McEvoy & Mahoney, 2012). IU has also been shown to have the strongest association with worry over and above other components in the IU model (positive beliefs about worry, negative problem orientation, and cognitive avoidance; Lachance, Ladouceur, & Dugas, 1999; Laugesen, Dugas, & Bukowski, 2003; Robichaud, Dugas, & Conway, 2003). Moreover, in a series of experimental studies, increasing levels of IU using different manipulation tasks led to higher levels of worry and distress (Buhr & Dugas, 2009; Ladouceur, Gosselin, & Dugas, 2000; Rosen & Knäuper, 2009). Although the vast

majority of research investigating the IU model has occurred outside the realm of psycho-oncology, preliminary evidence also indicates that IU is associated with worry in BCa survivors (Costa-Requena, Rodríguez, Fernández, Palomera, & Gil, 2011), and emotional distress in breast (Costa-Requena et al., 2011; Taha, Matheson, & Anisman, 2012), prostate (Eisenberg et al., 2015), and lung (Kurita, Garon, Stanton, & Meyerowitz, 2013) cancer survivors.

6.3 The Self-Regulatory Executive Function (S-REF) model

The S-REF model (Wells & Matthews, 1994, 1996) is a transdiagnostic model which explains the development and maintenance of emotional distress across a range of mental and physical health populations. According to the S-REF model, stored metacognitive beliefs (i.e. beliefs about thinking), opposed to cognitive beliefs, underlie RNT and persistent emotional distress.

The S-REF model posits that stored metacognitive beliefs activate a negative and sustained response style called the cognitive attentional syndrome (CAS). The CAS consists of RNT, fixating attention on potential signs of threat (e.g. constantly scanning for symptoms of BCa and being hypervigilant to negative BCa thoughts) and the use of maladaptive coping strategies (e.g. avoiding reminders of BCa, constantly seeking reassurance that BCa has not returned, and distracting attention away from negative thoughts about BCa). The CAS prolongs and intensifies distress via several pathways. For example, RNT increases the frequency and strength of negative thoughts; fixating attention on threat increases intrusive mental experiences and maintains the sense of threat; and the use of maladaptive coping strategies interferes with normal cognitive processes and prevents the opportunity to discover that cognitive beliefs are often erroneous. These strategies increase engagement in RNT and ultimately exacerbate emotional distress.

A wide range of metacognitive beliefs are specified in the S-REF model but five have been the primary focus of research: positive metacognitive, negative metacognitive beliefs, need to control thoughts, cognitive self-consciousness, and cognitive confidence. Positive metacognitive beliefs refer to beliefs about the benefits of or need to engage in RNT (e.g. “I must ruminate in order to find an answer to my sadness”). Negative metacognitive beliefs refer to beliefs about the uncontrollability and danger

of RNT (e.g. “my worrying is uncontrollable”). The need to control thoughts refers to the extent to which one believes that certain thoughts should be suppressed (e.g. “I need to control my thoughts at all times”). Cognitive self-consciousness refers to the tendency to self-monitor thoughts and focus attention inwards (e.g. “I pay close attention to the way my mind works and constantly examine my thoughts”). Cognitive confidence refers to the amount of confidence in one’s own cognitive functioning (e.g., “I constantly doubt my own memory”). These stored metacognitive beliefs predispose individuals to engage in CAS activity. Once activated, the CAS strengthens stored metacognitive beliefs by increasing accessibility to information that supports them. This further strengthens CAS activity such as RNT, which in-turn further exacerbates emotional distress.

6.3.1 Empirical support for the S-REF model

Metacognitive beliefs significantly predict emotional distress and RNT in the form of worry and rumination in a range of clinical and non-clinical populations (Wells, 2000, 2009; Wells & Fisher, 2015). Evidence also supports the central prediction of the S-REF model; that metacognitive beliefs give rise to RNT which in-turn leads to emotional distress (Wells, 2000, 2009; Wells & Fisher, 2015). While most empirical support for the model has occurred outside the realm of psycho-oncology, evidence of the model’s utility for understanding RNT and emotional distress in cancer is beginning to emerge. Metacognitive beliefs are associated with worry in breast and prostate cancer survivors (Cook et al., 2015a; Thewes, Bell, & Butow, 2013); and emotional distress in breast, prostate, colon (Cook et al., 2015a, 2015b; Quattropani, Lenzo, Mucciardi, & Toffle, 2015), and adolescent and young adult (Fisher et al., 2018) cancer survivors.

6.4 Summary

Both the IU model and the S-REF model suggest processes underlying RNT and emotional distress in BCa. Although both models overlap to some extent (i.e. positive beliefs about worry), the fundamental hypothesised causal psychological processes are distinct. As both models focus on a process underlying distress, they offer the potential for a transdiagnostic treatment approach, which may be more appropriate to BCa

patients who often present with mixed symptoms of anxiety, depression, trauma, and general distress.

6.5 Conclusion

There is preliminary evidence that both IU and metacognitive beliefs are associated with emotional distress in BCa survivors. However, the role of these constructs for explaining RNT has only partially been explored and no study to date has investigated the role of IU and metacognitive beliefs in BCa within the same study. Thus, it is unclear whether IU or metacognitive beliefs uniquely predict RNT and distress after controlling for each other. In the few studies in non-BCa populations investigating metacognitive beliefs and IU concurrently, metacognitive beliefs have consistently emerged as a stronger predictor of RNT and distress (Fergus & Wheless, 2018; Gerlach, Andor, & Patzelt, 2008; Khawaja & McMahon, 2011; Thielsch, Andor, & Ehring, 2015a, 2015b). It is not yet known whether the same finding would occur in BCa populations.

Moreover, all but one of the previous studies investigating metacognitive beliefs and IU in BCa has been cross-sectional, precluding the study of temporal precedence. Therefore, it is impossible to conclude whether IU and/or metacognitive beliefs are a cause, rather than a consequence of RNT and distress. In the one prospective study (Cook et al., 2015b), retrospective self-report measures were used to explore engagement in RNT and distress over the preceding week or month. However, such methods are often inaccurate due to recall biases (Fahrenberg, Myrtek, Pawlik, & Perrez, 2007; Hassan, 2006). For example, individuals are more likely to recall experiences that occurred more recently; and that are consistent with their current mood (Gorin & Stone, 2001; Hufford, Shiffman, Paty, & Stone, 2001). The risk of inaccuracy or bias is increased because emotional distress and engagement in RNT fluctuate over short intervals such as hours and days (Moberly & Watkins, 2008).

To overcome the difficulties associated with traditional retrospective self-report measures and to take account of short-term variability in RNT and distress in BCa, an alternative methodology that captures RNT and emotional distress shortly after being experienced is needed. One approach is experience sampling methodology (ESM), described in the following chapter.

Chapter 7. An Overview of Experience Sampling Methodology

7.1 Introduction

Most prospective research in psychology has focused on relationships between psychological constructs at two or three occasions often weeks or months apart. However, there is an increased awareness that psychological models of emotional distress are essentially dynamic (Bentall, 2004) leading to large variation in psychological constructs over short intervals. A growing number of researchers have therefore advocated for an alternative methodology that accounts for variability over short periods of time (Csikszentmihalyi & Larson, 2014). A method known as experience sampling methodology (ESM) was therefore devised. ESM, also referred to as ecological momentary assessment (EMA), is a longitudinal research method that involves asking participants to complete a short assessment about their current or recent experiences several times each day in everyday settings (Csikszentmihalyi & Larson, 1987; Stone & Shiffman, 1994). ESM has been used in numerous mental and physical health populations in both children and adults (Broen et al., 2016; Csikszentmihalyi & Larson, 2014; Curran, Beacham, & Andrykowski, 2004; Mujagic et al., 2015; Myin-Germeys et al., 2009).

7.2 Advantages of ESM

As ESM enables individual's current or recent experiences to be continuously captured as they fluctuate in everyday settings, it provides a rich ecological data set and enables dynamic relationships between variables to be evaluated in much more detail (Aldao, 2013; Gratz & Roemer, 2004; Myin-Germeys et al., 2009). It also minimises the likelihood of recall biases which are prevalent in traditional retrospective self-report measures (Dupuy, Beaudoin, Rhéaume, Ladouceur, & Dugas, 2001). For example, individuals are more likely to recall experiences that occurred more recently; and that are consistent with their current mood (Gorin & Stone, 2001; Hufford et al., 2001). Moreover, individuals experiencing psychological morbidity are more likely to recall emotionally negative rather than positive experiences during retrospective assessments (Fritzsche et al., 2010; Lepage, Sergerie, Pelletier, & Harvey, 2007).

7.3 Disadvantages of ESM

There are two main limitations of ESM: it can be perceived as intrusive and time-consuming by participants, leading to selection and sample biases (Cerin, Szabo, & Williams, 2001); and it increases the chance of participant reactivity due to multiple assessments and frequent observation. Participant reactivity refers to when participants' experiences or behaviors change due to their awareness of being observed (Parsons, 1974). Careful design and piloting of an ESM study is necessary to overcome these limitations.

7.4 Designing an ESM study

When designing an ESM study, the following need to be considered; sampling procedure, development of the ESM assessment, ESM equipment, recruitment process, and piloting of the ESM study.

7.4.1 Sampling procedure

Two main ESM sampling procedures exist: event-contingent sampling and signal-contingent sampling. In event-contingent sampling, participants are prompted to complete an assessment following every occurrence of an event of interest. This type of sampling is useful when the experience of interest revolves around a certain event (e.g. smoking). In signal-contingent sampling, participants are prompted to complete assessments at certain times each day. As this procedure provides a broader assessment of daily life, it is the preferred procedure in ESM studies (Palmier-Claus et al., 2011). Signal-contingent sampling can be split into two further categories: fixed and random sampling. Fixed sampling involves prompting participants to complete assessments at the same times each day. With fixed sampling, it is easier to conduct time-related statistical analysis. However, it has the potential for participants to become aware of when they will be prompted, potentially resulting in participants changing their routine in anticipation of an upcoming prompt or thinking about their responses prior to being prompted. Random sampling, on the other hand, involves prompting participants to complete assessments at random times each day. This has the advantage of preventing participants from being prepared for upcoming assessments. However, it may result in long periods of time with no assessment followed by numerous assessments in quick

succession, representing only a small part of daily variance. This issue can be resolved by using pseudo-random sampling, in which participants are prompted to complete assessments at random points within fixed time intervals (e.g. a random prompt within every 2-hour block of time).

It is also important to consider how often and for how long participants will be prompted to complete ESM assessments. This largely depends on the amount of data required to achieve a representative picture of the constructs being investigated. More prompts (and therefore more assessments) do not necessarily result in more data, as an overload of prompts may result in less compliance and higher dropout rates (Palmier-Claus et al., 2011). There is no "gold standard" for the frequency and length of time to prompt participants, However, ESM studies have ranged from 1-14 prompts per day over a period of 1-90 days (Exler et al., 2017; Kimhy et al., 2006; Sullivan, Khondkaryan, Dos Santos, & Peters, 2011; Wijesekera et al., 2018).

7.4.2 Development of the ESM assessment

The central component of any ESM study is the questions included in the ESM assessment (Palmier-Claus et al., 2011). A typical ESM assessment contains a collection of questions gathering information about participant's current or recent thoughts, feelings or experiences using simple, unambiguous questions such as "*how happy do you feel right now?", "just before the beep, how much were you worrying?"* or *"since the last beep, how sad have you felt?"*. As the aim is to capture current or recent experiences opposed to retrospective experiences, the wording of questions in ESM assessments fundamentally differ from those used in standard retrospective questionnaires. The development of an ESM assessment is often guided by global questionnaires but worded differently to be applicable in everyday life. For example, the following question from the HADS *"I feel tense or wound up"* could be changed to *"since the last beep, I have felt tense or wound up"*. The time taken to complete an ESM assessment should not exceed 2–3 minutes (Myin-Germeys et al., 2018). Responses to questions can be obtained using open-ended formats or numerical scales. While open-ended responses can be informative, they take considerably longer to complete. Therefore, ESM responses are typically obtained using numerical scales.

7.4.3 Choosing equipment: Paper & pen vs electronic methods

ESM was originally administered using paper-pen methods, which involves asking participants to carry around a paper-based pack of ESM assessments, and a signalling device with pre-programmed 'beeps' (usually a wristwatch). Participants are typically asked to complete a paper-based assessment when prompted by the preprogramed 'beep' on a wristwatch (paper-based ESM). However, paper-based ESM cannot ascertain that participants complete assessments when prompted. Thus, participants may forward- or back-log their responses (e.g. a participant may have a busy day and forget to complete the ESM assessments when prompted and instead complete them all at the end of the day), invalidating the aim of ESM to capture current or recent experiences as they fluctuate over time. Furthermore, carrying around a collection of paper-based ESM assessments is often made obvious to others, inviting questions and adding to participant burden (Thomas & Azmitia, 2016).

Electronic methods of administering ESM have therefore been developed, with the use of palmtop computers becoming extremely popular. The main advantage of 'palmtop ESM' is that assessments can be accurately 'time-stamped', making it possible to know when each assessment is completed. Thus, assessments not completed in a permitted time-frame can be discarded, solving the issue of forward- or back-logging. Electronic methods also have higher compliance rates than paper-pen methods. Stone, Shiffman, Schwartz, Broderick, and Hufford (2002) conducted an ESM study comparing paper-pen and palmtop methods. To assess paper-pen compliance, ESM paper assessment packs were fitted with photo-sensors that recorded the opening of the booklet - a prerequisite to completing an ESM assessment. Patients were instructed to only complete ESM assessments within 30 minutes of the prompt. The palmtop ESM software was programmed to only enable participants to complete assessments within this 30-minute window, ensuring that only relevant assessments were completed. Stone et al. (2002) found that participants completed the majority of ESM assessments using both the paper-pen (90%) and palmtop (94%) method. However, when compliance with the paper-pen method was verified by the photo-sensor, completion was much lower (11%).

While palmtop ESM is an improvement on paper-based ESM, palmtops are expensive and largely outdated making it difficult for researchers to find appropriate devices (Burgin, Silvia, Eddington, & Kwapil, 2013). Moreover, they still have to be carried around, adding to participant burden. Consequently, numerous smartphone applications (apps) have been developed for use in ESM studies (e.g. ClinTouch; MovisensXS; P.A.C.O; PIEL survey; metric-wire) and are quickly becoming the most preferred ESM data-collection method (Berkel, Ferreira, & Kostakos, 2017). This 'smartphone ESM' approach holds all the advantages of palmtop ESM methods but adds no physical burden and is a cheaper option.

7.4.4 Pilot testing of ESM

Researchers are strongly advised to pilot ESM studies before starting recruitment (Palmier-Claus et al., 2011), with Hektner, Schmidt, and Csikszentmihalyi (2007, p. 52) stating that *"it would be difficult to overstate the importance of pilot testing when doing an ESM study".* Piloting is not only important to ensure that the newly developed ESM assessments are understandable but also to ensure that the prompting schedule chosen is not too invasive and that the ESM equipment works without any errors (e.g. to check prompts are 'released' in the correct time frames or to check the prompts are audible to patients). Ideally, ESM studies should initially be piloted on a member of the research team followed by a small number of individuals who are in close contact with the chosen population (Hektner et al., 2007).

7.4.5 Recruiting participants into an ESM study

Considering the perceived invasive nature of ESM, recruiting participants into ESM studies can be difficult. ESM studies often have high participant dropout rates with only the most motivated taking part, leading to potential sample biases (Cerin et al., 2001). Despite high dropout rates, ESM studies have been able to recruit sample sizes ranging from 5–145 participants (Delle Fave & Massimini, 2004; Ruscio et al., 2015). To increase participant recruitment and compliance, it is highly recommended that, upon recruitment to the study, the researcher meets with the participants for an orientation session (typically lasting around 45 minutes to an hour; Hektner et al., 2007). During this session, the researcher should explain the ESM procedure to the participant, ensure that the participant understands all the questions in the ESM

assessment, train the participant to correctly use the ESM equipment, and allow the participant to complete some practice ESM assessments.

7.5 Data analysis in ESM

An important assumption of many statistical analyses is that data points are independent of one another (i.e. two individual data points are assumed not to be related to one another). In reality, this assumption is often violated. For example, when examining the effects of revision on test results, you might expect students from the same school to have similar results. This type of data is referred to as "nested" data. Not controlling for "nested" data can significantly bias results (Goldberg, 1992). As several measurements are taken from each participant over several days, ESM, by convention, involves "nested" data. Measurements are likely to be more comparable within-person opposed to between-person. Measurements may also be more comparable within days opposed to between days. Thus, ESM is nested at three levels: Level 1 (the lowest level) represents each individual ESM assessment; level 2 represents each day; and level 3 represents each participant. This multilevel data structure is presented in Figure 7.1. To deal with this multilevel structure, ESM data can be aggregated for each participant. This involves combining participants' individual questionnaire responses to give an overall score. However, this removes the multi-level structure of the ESM data, thus limiting the potential to explore within-person variability (one of the main advantages of ESM). Alternatively, a technique known as "multilevel modelling" (MLM) can be used. MLM enables the estimation of the amount of variation at each level and maintains the multilevel structure of the data. Considering, the nested nature of ESM, small sample sizes still result in a very rich dataset. For example, if 20 participants completed 10 questionnaires a day for six days, it would result in 1,200 data points (20 participants x 10 questionnaires x 6 days).

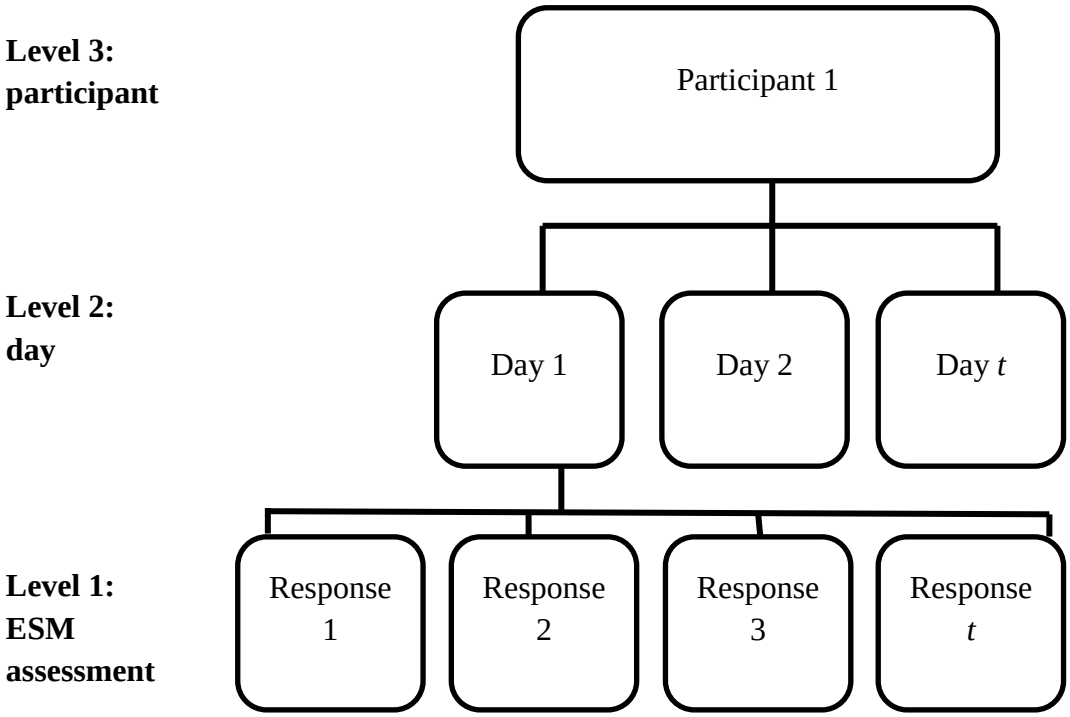


Figure 7.1: Multilevel structure of ESM data

7.6 Summary

This chapter has presented an overview of important practical issues that need to be considered when designing and implementing an ESM study. Taking these considerations into account and considering the advantages of smartphone ESM, the following chapter will describe a smartphone ESM study conducted to explore the role of IU and metacognitive beliefs model in the development and maintenance of RNT and emotional distress in BCa survivors. The feasibility of conducting a smartphone ESM study in BCa survivors was also tested.

Chapter 8. Study 4: An Experience Sampling Methodology Study of Emotional Distress in Breast Cancer Survivors - The Role of Metacognitive Beliefs and Intolerance of Uncertainty

8.1 Introduction

There is preliminary evidence that both IU and metacognitive beliefs are associated with emotional distress in BCa survivors. However, as highlighted in chapter 6, the role of these constructs for explaining RNT has only partially been explored and no study to date has investigated the role of IU and metacognitive beliefs in BCa within the same study. Moreover, all previous studies have relied solely on traditional retrospective self-report measures, which are often inaccurate due to recall biases (Fahrenberg, Myrtek, Pawlik, & Perrez, 2007; Hassan, 2006). Smartphone ESM can overcome these limitations.

Smartphone ESM has been used in numerous physical and mental health populations (Band, Barrowclough, Caldwell, Emsley, & Wearden, 2017; Palmier-Claus et al., 2012; Palmier-Claus et al., 2014; Seidel et al., 2016; van der Velden, Mulders, Drukker, Kuijf, & Leentjens, 2018; Westermann et al., 2017). Yet, only one smartphone ESM study has been conducted in BCa (focusing specifically on sleep disturbance and mood in patients receiving chemotherapy; Min et al., 2014), but none have been conducted in BCa survivors.

Therefore, the aims of this study were two-fold. The first was to assess the feasibility of using smartphone ESM in BCa survivors. This aim was exploratory and intended to help in designing future ESM studies. Specifically, we examined how representative recruited patients were of the study population, the degree of attrition, and the extent of compliance with the requirements of the study. The second aim was, using smartphone ESM, to explore whether IU and/or metacognitive beliefs uniquely predict anxiety, depression and RNT. As baseline anxiety, depression and RNT predict future anxiety, depression and RNT, respectively (Cook, Salmon, Hayes, Byrne, & Fisher, 2018; Thielsch et al., 2015a; Weber & Exner, 2013) baseline measures of these constructs, along with demographic variables, were included as covariates. Specifically, we hypothesised that: 1) IU and metacognitive beliefs would predict anxiety, depression and RNT; 2) IU and metacognitive beliefs would be better predictors of anxiety, depression and RNT than known covariates; and 3) metacognitive beliefs would predict anxiety, depression and RNT over and above IU.

8.2 Method

8.2.1 Design

We used a prospective cohort design. Independent variables were metacognitive beliefs and IU; covariates were demographic variables and baseline anxiety, depression and RNT; dependent variables were anxiety, depression, and RNT. Independent variables and relevant covariates were assessed at baseline using traditional self-report measures. Dependent variables were assessed using ESM. There is little theoretical guidance on the appropriate frequency and duration of an ESM study. Therefore, based on previous ESM studies (Marco & Suls, 1993; Myin-Germeys, Krabbendam, Jolles, Delespaul, & van Os, 2002), ESM assessments occurred six times daily for eight consecutive days.

8.2.2 Participants

Inclusion criteria were: diagnosis of BCa, completion of adjuvant therapy (excluding hormone replacement therapy), access to a smartphone with an android or IOS operating system, aged 18 years or older, and ability to speak and understand English. Patients considered by the clinical team or researcher to be too distressed or confused to give informed consent were excluded. The study was approved by the NHS North West Greater Manchester East ethics committee (15/NW/0925).

8.2.3 ESM protocol

ESM assessments were delivered by an app downloaded onto participants' smartphones. The Participation in Everyday Life Survey application (PIEL app; www.pielsurvey.org) was used for iOS devices and the MovisensXs app (www.movisens.com) was used for android devices. Participants were 'prompted' to complete each ESM assessment by a pre-programmed 'beep' within the app. To capture sufficient variability in daily experiences, prompts were delivered at pseudo-randomised intervals with a minimum of 60 minutes between each prompt. If an assessment was not completed within 15 minutes of the prompt, it was no longer accessible. The time-frame during which participants were prompted was individualised by programming the app to only prompt participants during their average waking hours.

8.2.4 Procedure

Patients were recruited through a post-treatment BCa clinic at a National Health Service (NHS) teaching hospital in North-West England between September 2016 and July 2017. Clinical staff identified potentially eligible patients from their medical records and sent them a recruitment letter and information sheet which they received, along with their appointment letter, 3-4 months prior to their appointment. Immediately following their appointment, patients were asked by the BCa nurse conducting the appointment whether they had access to a smartphone, and if so, would they like to speak to the researcher (JT) to learn more about the study. Those who agreed met with the researcher who answered questions, screened for eligibility, and obtained written informed consent from those who agreed to participate. Participants then completed the baseline questionnaires in the clinic. Following this, the ESM app was downloaded onto each participant's smartphone and programmed to begin on an agreed date within the following seven days. The ESM assessments continued for eight consecutive days. On the second day of the ESM schedule, participants were contacted by the researcher (JT) to ensure that the app was functional and that they still wanted to participate.

8.2.5 Measures

8.2.5.1 Independent variables

Metacognitive beliefs were assessed using the metacognitions questionnaire-30 (MCQ-30; Wells & Cartwright-Hatton, 2004). The scale contains 30 items measuring five metacognitive beliefs: positive metacognitive beliefs about the benefits of or need to engage in RNT; negative metacognitive beliefs about the uncontrollability and danger of RNT; the need to control thoughts; cognitive self-consciousness; and lack of cognitive confidence. Each subscale yields a total score ranging from 6 to 24 with higher scores indicating greater conviction in metacognitive beliefs. The MCQ-30 has been validated for use with cancer patients (Cook, Salmon, Dunn, & Fisher, 2014). In the current sample, Cronbach's α of MCQ-30 subscales ranged from 0.73 ('need to control thoughts') to 0.9 ('cognitive confidence') indicating adequate to excellent internal consistency.

IU was assessed using the Intolerance of Uncertainty Scale (IUS; Buhr & Dugas, 2002). The IUS is a 27-item self-report questionnaire measuring ones' emotional, cognitive and behavioural reactions to uncertain situations. Total scores range from 27 to 135 with higher scoring indicating greater IU. The scale has been validated in BCa (Costa-Requena et al., 2011) and has excellent internal consistency in the current sample (Cronbach's α: 0.96).

8.2.5.2 Covariates

Demographic and clinical data was collected by self-report. Data collected were age at diagnosis, tumour stage at diagnosis, employment status, living alone or not, and time since finishing adjuvant therapy.

Baseline anxiety and depression were assessed using the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983). The HADS is a 14-item self-report questionnaire measuring anxiety (7 items) and depression (7 items). Each subscale is scored from 0 to 21, with higher scores indicating greater anxiety or depression. The HADS is one of the most well-validated outcome measures of anxiety and depression in BCa populations (Hann et al., 1999; Hashim, 2016; Villoria & Lara, 2018). In the current sample, both subscales have good internal consistency (Cronbach's α: 0.88 for anxiety; 0.81 for depression).

Because no RNT scale has yet been validated in BCa or a similar population, baseline RNT was assessed as two separate variables: worry and rumination. Worry was assessed using the Penn State Worry Questionnaire (PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990), a 16-item self-report questionnaire measuring the intensity and excessiveness of worry independent of worry content. The scale yields a total score of 16 to 80 with higher scores indicating greater worry. The PSWQ has good internal consistency in cancer populations (e.g. Lehto & Cimprich, 2009) and has good test-retest reliability (Meyer et al., 1990). The scale has excellent internal consistency in the current sample (Cronbach's α: 0.93). Rumination was assessed using the ruminative response scale (RRS; Nolen-Hoeksema, 1991), a 22-item self-report scale. Scores range from 22 to 88 with higher scores indicating higher levels of rumination. The scale has good internal consistency in cancer populations (e.g. Schellekens et al.,

2017) and has excellent internal consistency in the current sample (Cronbach's α: 0.95).

8.2.5.3 Dependent variables (ESM assessment)

The ESM assessment included three questions – one for each dependent variable; anxiety, depression, and RNT (see Figure 8.1). To limit potential recall biases and capture data representative of participant's daily experiences, questions in the ESM assessment began with the phrase *"since the last beep".* The questions relating to anxiety and depression were similar to those used in previous ESM assessments (Bylsma, Taylor-Clift, & Rottenberg, 2011; Hartley, Haddock, Vasconcelos e Sa, Emsley, & Barrowclough, 2014; Moberly & Watkins, 2008; Peeters, Nicolson, Berkhof, Delespaul, & deVries, 2003). The question relating to RNT was adapted from the Cognitive Attentional Syndrome Scale (CAS-I; Wells, 2009), a retrospective self-report measure that assesses, along with other constructs, RNT.

8.2.6 Patient and public involvement

Feedback regarding the ESM design was informally obtained from a group of BCa survivors who were members of a local cancer support group. This feedback led to significant revisions of the ESM assessment, including rephrasing questions so that they were preceded by "since the last beep" opposed to "just before the beep" as patients felt it was difficult to pin-point their exact feelings immediately prior to the 'beep'. The number of questions in the ESM assessment was also reduced as participants felt it was originally too burdensome.

8.2.7 Pilot testing

As described in chapter 7, pilot testing newly developed ESM assessments and software is strongly advised (Palmier-Claus et al., 2011). Thus, the ESM procedure was informally piloted on a member of the research team. Technical issues relating to the delivery of the ESM assessment on the smartphone app were resolved accordingly (i.e. the sampling schedule initially reset every time the smartphone was turned off). Following this, the ESM procedure was informally piloted on six BCa nurses at the participating NHS hospital. Further technical issues were dealt with (i.e. initially, the volume of 'prompts' could not be muted, causing unwanted disruption during the

nurses' schedules). Moreover, it indicated that the average time it took them to complete the ESM questionnaire was 90 seconds, falling within the 2-3 minute range recommended (Palmier-Claus et al., 2011).

8.2.8 Feasibility

8.2.8.1 Attrition

We calculated the proportion of patients who agreed to meet with the researcher, consented to take part, and dropped out after giving consent.

8.2.8.2 Sample representativeness

We compared the education level, ethnicity, marital status, and employment status of our sample with a sample of cancer patients (n=229) recruited from the same hospital, who took part in a recent cross-sectional study exploring worry and distress using traditional retrospective self-report measures (Cook et al., 2015a).

One potential barrier to using smartphone ESM in BCa survivors is smartphone accessibility. Most BCa patients (80%) are aged 50 years or older, with a median age of 62 (Horner et al., 2009). However, only 55% of people aged between 55-64, and only 18% of people aged 65 or older own a smartphone in the UK (Ofcom, 2018). Thus, to explore representativeness of the sample age, the median age of our sample was compared to the median age of BCa patients according to statistics from the National Cancer Institute (Horner et al., 2009).

8.2.8.3 Compliance

We calculated the number of ESM assessments completed per participant. In line with previous ESM studies (e.g. Hare, Gracey, & Wood, 2016; Moberly & Watkins, 2008; Myin-Germeys, van Os, Schwartz, Stone, & Delespaul, 2001), minimum compliance was defined as completion of at least one third of assessments. Level of compliance was also compared to that in a smartphone ESM study that used a similar ESM schedule (i.e. four times a day for seven consecutive days) and indicated response rates in samples from four distinct populations: schizophrenia, substance dependence, anxiety disorders, and a non-clinical sample (Johnson et al., 2009).

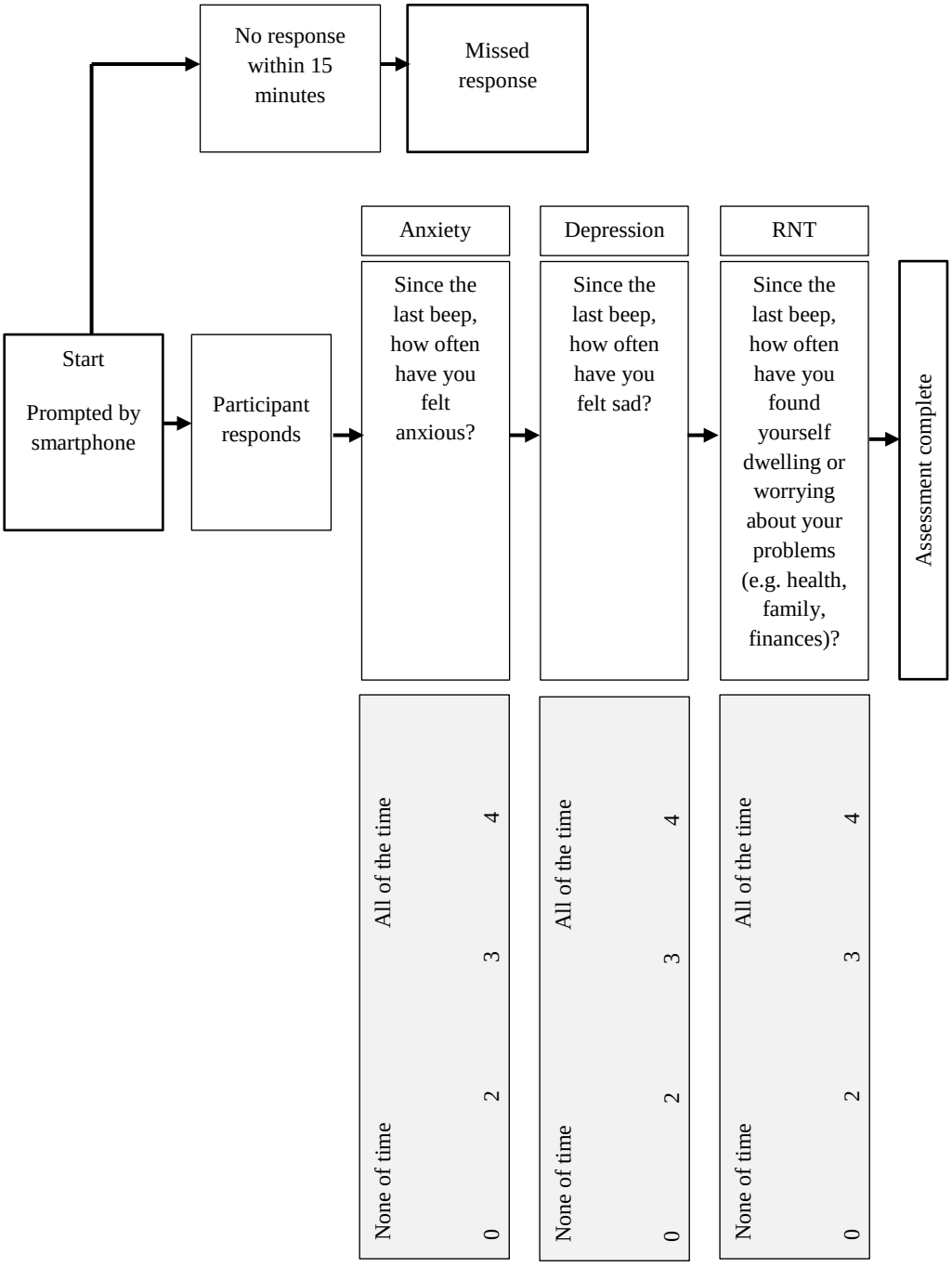


Figure 8.1: Summary of ESM procedure and details of ESM assessment

8.2.9 Statistical analysis

8.2.9.1 Hypothesis 1 & 2

Data obtained by ESM is nested within three levels: assessments (level 1) nested within days (level 2) nested within individuals (level 3). Therefore, multilevel models of longitudinal data were used to test the hypotheses. Multilevel models can include both fixed effects (which take the same value across all individuals, days and assessments) and random effects which allow scores to differ between individuals, days, and assessments). Days were coded sequentially (e.g. 1 represents day 1) and assessments were coded in hours (e.g. 9.5 represents 9.30am). To improve the chance of a well-specified multilevel model, Diggle et al. (2002) recommend a procedure that firstly establishes an appropriate random effects structure for the model, then establishes appropriate fixed effects to include. This procedure was conducted separately for each dependent variable (i.e. anxiety, depression and RNT). Firstly, a model containing all possible fixed effects (i.e. independent variables and covariates) was fitted (excluding interactions between fixed effects) using restricted maximum likelihood techniques. Various models with different random effects structures were compared, from inclusion of no random effects, random intercepts at either the individual or day level, or both, and the addition of a random assessment time term at either the day level (to explain differences in responses to assessments *within* individuals over the course of a day; e.g. patient 1 might have responses that fall more quickly over the course of day 1 than day 2) and at the individual level (to explain differences in responses to assessments *between* individuals over the course of a day i.e. patient 1 might have responses that fall more quicker over certain days than patient 2), or both. The model reporting the lowest Akaike's Information Criterion (AIC; Akaike, 1974) was selected as the model of best fit. Fixed effects were then examined and sequentially removed in a backwards selection process using likelihood ratio tests. Fixed effects were removed from the model in order of least significant, until all terms remaining in the model were significant. P values of less than 0.05 were used to measure significance of the overall model. However, 95% confidence intervals (CIs) were used to measure the significance of individual fixed effects as p values for individual fixed effects in longitudinal multi-level models are not recommended (Bates, Mächler, Bolker, & Walker, 2015). Fixed effects were identified as non-

significant if their 95% CIs contained zero. Fixed terms for intercept and assessment time (i.e. timing of each individual assessment) were always included in the model. 95% CIs were calculated using bootstrapping procedures based on 500 bootstraps. Fixed effects examined for inclusion in each model are presented in Table 8.1.

8.2.9.2 Hypothesis 3

To explore whether metacognitive beliefs predict anxiety, depression and RNT over and above IU, IU was controlled for in the model (i.e. the model of best fit). However, to be able to control for IU in the model, IU has to be present. Thus, in situations where IU was not present in the model, the analysis was re-ran and IU was 'forced' into the model to allow it to be controlled.

8.2.10 Missing data

In line with previous ESM studies (Hare et al., 2016; Moberly & Watkins, 2008; Myin-Germeys et al., 2001), participants who completed less than a third of assessments were excluded from the analysis. Missing data for baseline measures (i.e. independent variables and covariates) were imputed using the 'MICE' package in R (Zhang, 2016). All analyses were performed in R [3] version 3.5.0, using packages lme4 (Bates et al., 2015).

8.3 Results

8.3.1 Sample characteristics

Forty-five of the 51 patients who took part in the study completed at least a third of the ESM assessments and were included in the analysis. There were no significant differences between completers (n=45) and non-completers (n=6) on clinical or demographic characteristics. The baseline clinical and demographic characteristics of the final sample (n=45) are shown in Table 8.2.

Table 8.1: Fixed effects entered into multilevel models for each dependent variable

Fixed effects (independent variables & covariates)	Dependent variables		
	Anxiety	**Depression**	**RNT**
(Intercept)	✓	✓	✓
Assessment start time	✓	✓	
Anxiety	✓		
Depression		✓	
Worry			✓
Rumination			✓
Intolerance of uncertainty	✓	✓	✓
Positive metacognitive beliefs about the benefits of or need to engage in RNT	✓	✓	✓
Negative metacognitive beliefs about the uncontrollability and danger of RNT	✓	✓	✓
Cognitive confidence	✓	✓	✓
Need to control thoughts	✓	✓	✓
Cognitive self-consciousness	✓	✓	✓
Age	✓	✓	✓
Employment status[a]	✓	✓	✓
Living alone or not	✓	✓	✓
Tumour stage at diagnosis[b]	✓	✓	✓
Time since finishing primary adjuvant therapy	✓	✓	✓

Note. RNT = repetitive negative thinking [a] coded as employed vs. unemployed; [b] coded as non-metastatic vs. metastatic

8.3.2 Feasibility

8.3.2.1 Attrition

Of the 189 patients invited to participate, 81 (37.6%) agreed to meet with the researcher, 51 (26.9% of those approached and 62.9% of those who met with the researcher) consented to take part and completed baseline questionnaires, and none dropped out after consenting.

Sample representativeness

Most participants were White Caucasian (84.4%), had a school qualification or higher (91.1%), were married, in a civil partnership or cohabiting (73.3%), and were employed (66.7%). This compared to the comparator sample (Cook et al., 2015a) as follows: 98% White Caucasian, 58% had a school qualification or higher, 66% married, in a civil partnership or cohabiting, and 38% employed. The median age of our sample was 58 (range 35-81), slightly lower than the median age of 62 of the general BCa population (Horner et al., 2009).

8.3.2.2 Compliance

The compliance rate was high (88.2% of patients). After removing patients who did not meet the compliance criterion, mean compliance rates were 64.8%, (SD=16%). This compared to mean compliance rates of 83% (SD=16%) in non-clinical patients, 80% (SD=18%) in patients with substance dependence, 73% (SD=18%) in patients with anxiety disorders, and 69% (SD=16%) in patients with schizophrenia (Johnson et al., 2009).

Table 8.2: Clinical and demographic sample characteristics at baseline

Total number of participants	45
Median age [min, max]	58 [35,81]
Median age at diagnosis [min,max]	56 [30,79]
Median Days since finishing primary adjuvant therapy [min,max]	127 [9,1538]
Median baseline anxiety [min,max]	6.0 [0,20]
Median baseline depression [min,max]	3.0 [0,18]
Gender (%)	
Male	1 (2.2)
Female	44 (97.8)
Employment (%)	
Employed	30 (66.7)
Unemployed	15 (33.3)
Living Arrangements (%)	
Living alone	7 (15.6)
Not living alone	37 (82.2)
Highest Qualification (%)	
None	4 (8.9)
School qualification or higher	41 (91.1)
Marital Status (%)	
Married/civil partnership/cohabiting	33 (73.3)
Single/divorced/separated/widowed	12 (26.7)
Current BCa episode was a recurrence (%)	
Yes	4 (8.9)

No	40 (88.9)
Tumour Stage at diagnosis (%)	
Non-metastatic	42 (93.3)
Metastatic	1 (2.2)
Not reported	2 (4.4)
Adjuvant therapy (%)	
Chemotherapy	
Yes	17 (37.8)
No	27 (60.0)
Radiotherapy	
Yes	30 (66.7)
No	14 (31.1)
Ethnicity (%)	
White Caucasian	38 (84.4)
Other	7 (25.6)

8.3.3 Hypothesis 1 & 2

Anxiety. The model of best fit (i.e. the model reporting the lowest AIC) for anxiety included a random intercept and random assessment time term at the day level, and a random intercept and random assessment time term at the individual level. This indicates that there was significant variation in anxiety scores both within and between individuals over the course of each day. After accounting for the intercept and assessment time, baseline anxiety (ß=0.09) was the only variable to significantly predict anxiety (Table 8.3).

Depression. The model of best fit for depression included a random intercept and random assessment time term at the day level, and a random intercept and random

assessment time term at the individual level. This indicates that there was significant variation in depression scores both within and between individuals over the course of each day. After accounting for the intercept and assessment time, baseline depression (ß=0.09) was the only variable to significantly predict depression (Table 8.4).

RNT. The model of best fit for RNT included a random intercept and random assessment time term at the day level, and a random intercept and random assessment time term at the individual level. This indicates that there was significant variation in RNT scores both within and between individuals over the course of each day. After accounting for the intercept and assessment time, negative metacognitive beliefs about the uncontrollability and danger of RNT (ß=0.07) was the only variable to significantly predict RNT (Table 8.5).

8.3.4 Hypothesis 3

As negative metacognitive beliefs about the uncontrollability and danger of RNT but not IU predicted RNT, the analysis was re-ran forcing IU into the model. After accounting for the intercept and assessment time at the day level and individual level, negative metacognitive beliefs about the uncontrollability and danger of RNT significantly predicted RNT (ß=0.06) over and above IU. All other variables were non-significant predictors of RNT (Table 8.6).

Table 8.3: Final multilevel model of anxiety

Fixed effects	**ß (95% CI)**
(Intercept)	0.62 (0.28,0.95)*
Assessment start time	-0.03 (-0.04,-0.02)*
Baseline anxiety	0.09 (0.05,0.13)*

Note. ß = beta; CI = 95% confidence interval; *= significant i.e. 95% confidence intervals do not contain zero

Table 8.4: Final multilevel model of depression

Fixed effects	**ß (95% CI)**
(Intercept)	0.61 (0.34,0.87)*
Assessment start time	-0.02 (-0.04,-0.01)*
Baseline depression	0.09 (0.05,0.14)*

Note. ß = beta; CI = confidence interval; *= significant i.e. 95% confidence intervals do not contain zero

Table 8.5: Final multilevel model of repetitive negative thinking (RNT)

Fixed effects	**ß (95% CI)**
(Intercept)	0.61 (0.18,1.04)*
Assessment start time	-0.03 (-0.04,-0.02)*
Negative metacognitive beliefs about the uncontrollability and danger of RNT	0.07 (0.03,0.1)*

Note. RNT; repetitive negative thinking; ß = beta; CI = confidence interval; *= significant i.e. 95% confidence intervals do not contain zero

Table 8.6: Final multilevel model of repetitive negative thinking (RNT) with the forced entry of intolerance of uncertainty (IU)

Fixed effects	**ß (95% CI)**
(Intercept)	0.59 (0.14,1.04)*
Assessment start time	-0.03 (-0.04,-0.02)*
Intolerance of uncertainty	0.00 (-0.01. 0.01)
Negative metacognitive beliefs about the uncontrollability and danger of RNT	0.06 (0.0,0.12)*

Note. RNT; repetitive negative thinking; ß = beta; CI = confidence interval; *= significant i.e. 95% confidence intervals do not contain zero

8.4 Discussion

The first aim of this study was to examine the feasibility of conducting a smartphone ESM study in a population of BCa survivors. The findings are promising. Most patients complied with the ESM schedule, and mean compliance rates were comparable to those in other populations using smartphone ESM (Johnson et al., 2009). Moreover, the sample was representative of BCa patients with regards to most demographic variables. The mean age was similar to the general BCa population (Horner et al., 2009); and the ethnicity and marital status were similar to those in a previous study of cancer patients in the same hospital (Cook et al., 2015a). However, more patients were employed and had more years of education than those in the previous study (Cook et al., 2015a). While attrition before consent was modest, no attrition occurred after consent.

The second aim was to explore whether IU and/or metacognitive beliefs predict anxiety, depression and RNT in BCa survivors. Negative metacognitive beliefs about the uncontrollability and danger of RNT were a better predictor of RNT than IU and known covariates. Moreover, negative metacognitive about the uncontrollability and danger of RNT remained a significant predictor of RNT even after controlling for IU. This is in line with the central tenet of the S-REF model that proposes that, while all metacognitive beliefs are important in the development and maintenance of emotional distress, negative metacognitive beliefs about the uncontrollability and danger of RNT are of more direct importance and can be viewed as the 'turbocharger' behind emotional distress (Wells & Matthews, 1994, 1996). In clinical and non-clinical populations, negative metacognitive beliefs about the uncontrollability and danger of RNT have also emerged as the strongest predictor of RNT and account for additional variance in predicting RNT beyond the effects of IU (Fergus & Wheless, 2018; Gerlach et al., 2008; Khawaja & McMahon, 2011; Thielsch et al., 2015a, 2015b).

Surprisingly, however, neither metacognitive beliefs nor IU predicted anxiety or depression. The only variables to predict anxiety or depression were baseline anxiety and depression, respectively. This does not support the study hypotheses and is inconsistent with previous cross-sectional (Cook et al., 2015a; Costa-Requena et al., 2011; Quattropani et al., 2015; Taha et al., 2012) and prospective (Cook et al., 2015b)

findings in BCa using traditional retrospective self-report measures. One potential reason for this finding is that the theoretical explanations of the S-REF model and IU model are insufficient for understanding emotional distress in BCa survivors. As most previous studies in BCa have been cross-sectional, it may be that metacognitive beliefs and IU are a consequence rather than cause of emotional distress. However, this does not explain the findings of Cook et al. (2015b) in which metacognitive beliefs prospectively predicted distress in BCa. Instead, it is possible that the S-REF model applies to BCa survivors' long-term retrospective recall of distress, but not their real-life distress as explored in this study. However, this seems unlikely (for the S-REF model) because metacognitive beliefs predicted RNT, which has repeatedly been shown to underlie emotional distress (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & Van Ijzendoorn, 2007; Calmes & Roberts, 2007; Mogg & Bradley, 1998; Muris & van der Heiden, 2006; Wells, 2000).

Other potential reasons why neither metacognitive beliefs nor IU predicted anxiety or depression are suggested by the low levels of distress reported by the sample and the inclusion of baseline distress as a predictor variable. First, the low levels of distress reported by the sample may have resulted in insufficient variance in distress levels. Moreover, as the S-REF model and IU model were developed to explain distress in clinically distressed populations, IU and metacognitive beliefs should (according to theory) be stronger predictors of anxiety and depression in a more distressed sample. Second, because of the inevitable correlation between baseline distress and future distress, the variance in baseline distress that arises from metacognitive beliefs and IU is likely to account for most of the variance in future distress. Thus, controlling for baseline distress leaves minimal variance left to account for. This may mask the effects of important psychological processes that underlie distress, such as metacognitive beliefs and IU.

Several limitations of this study need to be considered. First, most patients reported low levels of anxiety and depression. Thus, findings should not be generalised to clinically distressed BCa patients without further investigation. Second, although ESM was used to capture recent experiences, the questions within the ESM assessment were still retrospective (i.e. "since the last beep"). Thus, while recall biases are less likely due to the shorter time between experiences and reporting of those experiences,

they cannot be excluded. Further studies should use purely momentary questions (i.e. "just before the beep"). Third, although dependent variables were assessed using ESM, independent variables were assessed using only traditional retrospective self-report questionnaires. Further studies should assess both dependent and independent variables using ESM to reduce recall biases and enable testing through lagged predictive analyses leading to more conclusive findings regarding temporal precedence.

In summary, although metacognitive beliefs and IU did not predict anxiety and depression, negative metacognitive beliefs about the uncontrollability and danger of RNT were the strongest predictor of RNT even after controlling for IU, partially supporting the applicability of the S-REF model in BCa survivors. However, to pursue the findings of this study and determine reasons for the failure of IU and metacognitive beliefs to predict emotional distress, replication in a more distressed sample is needed. Our results suggest that conducting a smartphone ESM study in a population of BCa survivors is feasible. Though, further exploration of differences in employment rates and level of education between BCa patients who take part in studies using retrospective methods and those using ESM would be beneficial. Overall, smartphone ESM provides a promising new direction to help understand the psychological processes involved in the development and maintenance of emotional distress in BCa survivors.

Chapter 9. General Discussion and Conclusions

9.1 Introduction

There were two broad aims of this thesis; first, to advance knowledge and understanding regarding the efficacy of psychological treatments for emotional distress in BCa; and second, to develop a better understanding of the psychological processes associated with the development and maintenance of emotional distress in BCa survivors. To achieve these aims, three main studies were conducted. Each will now be discussed in turn.

9.2 A Systematic Review of the Quality of Randomised Controlled Trials of Psychological Treatments for Emotional Distress in Breast Cancer

Previous meta-analyses conclude that efficacious psychological treatments for emotional distress in BCa exist. These conclusions inform health care policies and clinical practice guidelines internationally which specify that psychological treatments should be available to BCa patients as part of their routine care (Dauchy et al., 2012; Holland et al., 2011; Howell et al., 2009; Li et al., 2016; National Breast Cancer Centre, 2003; National Comprehensive Cancer Network, 2003; National Institute for Clinical Excellence, 2004; Page & Adler, 2008; Reese et al., 2017; Tit et al., 2017). However, confidence in these conclusions can only be confirmed if the RCTs they are based on are of good quality. Therefore, a systematic review of the methodological quality of RCTs of psychological treatments for emotional distress in BCa was conducted. Overall, methodological quality was low. Generic design elements were limited in most trials: only 15% specified as an inclusion criterion that participants were distressed; only 10% controlled for concomitant treatments; and only 11% reported the clinical significance of findings. Design elements specific to psychotherapy RCTs were also implemented poorly: only 51% used treatment manuals; only 8% used certified trained therapists; and monitoring of adherence and competence occurred in only 15% and 4%, respectively. Thus, the current view that efficacious psychological treatments exist for emotional distress in BCa patients is based on poor quality RCTs. If relevant health policies are to be adequately empirically informed, meta-analyses and future RCTs must account for important methodological issues presented in this review.

9.3 Do Manualised Psychological Treatments Alleviate Emotional Distress in Breast Cancer Patients? An Individual Patient Data Meta-Analysis

As no previous meta-analysis focused specifically on clinically distressed patients, excluded non-manualised treatments, or evaluated the clinical significance of treatments, the practical relevance of their findings is questionable. Consequently, an IPD-MA of RCTs of manualised psychological treatments for emotional distress in BCa was conducted. Treatment efficacy was evaluated using both effect size analysis and, for the first time in BCa, clinical significance analysis. Analyses were conducted for the total sample (including distressed and non-distressed patients), and the clinically distressed sub-sample. In the total sample, the two methods of analysis converged on a disappointing picture. Controlled effect sizes were non-significant; and the likelihood of improving was only 6-8% higher for treated than control patients at post-treatment, and no higher at follow-up. Findings remained disappointing in the distressed sub-sample. Despite benefits from both analyses at post-treatment, there were no benefits for treated patients relative to controls at follow-up. Moreover, the clinical significance analysis showed that post-treatment benefits were very small: only 28-32% of treated patients recovered compared to 17-27% of controls. These findings show that psychological treatments do not alleviate emotional distress for most BCa patients. Thus, the general conclusion that efficacious psychological treatments for emotional distress in BCa exist appears to be wrong. The clinical utility of implementing these treatments into routine practice is questionable. More efficacious psychological treatments are urgently needed for BCa patients with emotional distress.

9.4 An Experience Sampling Methodology Study of Emotional Distress in Breast Cancer Survivors - The Role of Metacognitive Beliefs and Intolerance of Uncertainty

Having identified the urgent need to develop more efficacious psychological treatments for BCa patients with emotional distress, the second aim of this thesis was to develop a better understanding of the psychological processes associated with the development and maintenance of emotional distress in BCa survivors. To do this, a smartphone ESM study explored whether the core components of the S-REF model

(metacognitive beliefs) and/or the IU model (intolerance of uncertainty) predicted RNT and emotional distress in BCa survivors. As this was the first smartphone ESM study in BCa survivors, the feasibility of conducting a smartphone ESM study in this population was also tested. In line with the S-REF model, negative metacognitive beliefs about the uncontrollability and danger of RNT were the strongest predictor of RNT and were a better predictor of RNT than IU and known covariates. Moreover, negative metacognitive beliefs about the uncontrollability and danger of RNT remained a significant predictor of RNT even after controlling for IU. Unexpectedly, neither metacognitive beliefs nor IU predicted anxiety or depression. This is incongruent with the central prediction of both the S-REF model and IU model. This may suggest that the S-REF and IU model are insufficient for understanding emotional distress in BCa survivors. However, due to the low levels of distress reported by the sample, further investigation is needed to clarify this. The low rate of attrition, the high compliance, and the representativeness of the sample in this study indicate that conducting a smartphone ESM study in BCa survivors is highly feasible. This provides an alternative and promising new approach of understanding emotional distress in BCa survivors.

9.5 Limitations of the thesis

While limitations of the included studies have been documented in their respective chapters, some deserve further attention.

9.5.1 The quality scale used in study 2

The first issue relates to the quality scale used in study 2. Although the POMRF has been used to assess the quality of psychotherapy RCTs in numerous mental and physical health populations (Arnberg & Öst, 2014; Öst, 2008, 2014; Öst et al., 2015; Öst & Ollendick, 2017; Öst et al., 2016; Sloan et al., 2017; Swain et al., 2015; Swain et al., 2013), little is known about the scale's psychometric properties. Ensuring a quality scale is valid and reliable minimises the likelihood of errors when determining the quality of a scientific literature (Olivo et al., 2008). While the inter-rater reliability of the POMRF has been confirmed in this thesis and by others (Öst, 2008), the criterion, content, face, and construct validity of the POMRF has not. This is a limitation of most quality tools (Olivo et al., 2008). As the scale was developed by a

group of experts in psychotherapy outcome research, one would assume the tool would have reasonable face and content validity. Nonetheless, this tool needs to be validated to increase confidence in conclusions regarding trial quality.

Another issue relates to the items included in the POMRF. As discussed in chapter 2, methodological quality is a multidimensional construct. Thus, there will always be differing opinions on which items should be included in a quality scale. Any conclusions made based on the findings from study 2 can only relate to the aspects of quality assessed by the POMRF.

Finally, many items in the POMRF are ambiguous in their scoring criteria. This opens the possibility to differing opinions of what is required for a specific score. Although inter-rater reliability between reviewers was high, standardisation of scoring, for example, devising a manual with examples for each item, analogous to the "explanation and elaboration" document provided by CONSORT (Boutron, Moher, Altman, Schulz, & Ravaud, 2008) would likely further help this situation.

9.5.2 Cut-off points used to define functionality in study 3

An important limitation of study 3 relates to the calculation of the cut-off point used to define patients as part of a functional population post-treatment. As described in chapter 4, when using the Jacobson clinical significance method, there are three available methods for defining this cut-off point. Cut-off point (c) is the least arbitrary method as it is based on the relative probability of a patient's post treatment score belonging to either the functional or dysfunctional population. However, as appropriate normative data on functional populations were not available on the outcome measures used in study 3, only cut-off point (a) could be used (i.e. patient's post-treatment score falls outside the range of dysfunctionality, defined as falling two or more SDs beyond the mean of the dysfunctional population, in the direction of functionality). As described in chapter 4, cut-off point (a) may under- or over-estimate the score required for a patient to be classed as part of the functional population depending on the overlapping distribution between functional and dysfunctional populations. If the functional and dysfunctional distributions overlap, cut-off point (a) will be conservative. If they do not overlap, cut-off point (a) will be lenient. Without access to normative functional data, the extent of this overlap remains unknown.

9.5.3 Generalisability of findings from study 3

Another issue relates to the RCTs included in study 3. As only 17 of the 26 eligible RCTs provided IPD, not all eligible trials were included in the analysis, potentially compromising the generalisability of the results. However, no difference in controlled effect sizes were found between the 17 included RCTs and the 9 not included, suggesting that the included RCTs are representative of eligible published trials. Furthermore, to allow comparisons of clinically significant change between RCTs and provide greater confidence in results, only RCTs using the HADS, CES-D or POMS, the most widely used (Temple et al., 2018) and well-validated (Hann et al., 1999; Johnston et al., 2000; Nyenhuis et al., 1999) outcome measures in BCa, were included. This could also potentially compromise the generalisability of results, as not all RCTs evaluating manualised psychological treatments were included.

In contrast, although inclusion was restricted to RCTs using the HADS, CESD, or POMS, outcome data measured using different relevant measurement tools (i.e. HADS-D & CES-D for depression; and HADS-T & POMS-TMD for general distress) were still included. Not all measures are equal in their psychometric properties, including responsiveness to change. Therefore, the proportion of patients classed as having achieved reliable or clinically significant change may differ depending on the outcome measure used.

9.5.4 Clinical diversity of findings from study 3

RCTs included in study 3 were diverse with regards to at what point in the disease trajectory patients received psychological treatment (i.e. soon after diagnosis, during medical treatment, and survivorship). Treatments have been shown to work differently depending on the point in the disease trajectory that patients receive them (Zimmermann et al., 2007). PTSD literature has also suggested that early intervention can do more harm than good (Litz et al., 2002; Rose et al., 2003). Therefore, it may not be appropriate to aggregate across patients receiving treatment at different points in the disease trajectory. Unfortunately, when requesting IPD, participants' trajectory stage was not requested, and this information could not be extracted from published reports as most RCTs included patients at different points in the disease trajectory. Therefore, the influence of trajectory stage on outcome could not be explored.

However, for most outcomes, heterogeneity between trials was low suggesting that trajectory stage did not have much influence on treatment outcome.

Included RCTs were also diverse with regards to the type of psychological treatment used. To account for this, the influence of treatment type on outcome was investigated. However, as most treatments were broadly defined and involved a combination of two or more approaches, treatment groups had to be categorised using broad definitions. Grouping treatments that have important differences in rationale and procedures might obscure differences between treatments. For example, at post-treatment, the overall RD for recovery in depression for CBT was 0.02. However, RDs for individual CBT trials ranged from -0.08 (Groarke, Curtis, & Kerin, 2013) to 0.33 (Savard, Simard, Ivers, & Morin, 2005). To account for this variability, findings for each trial are presented individually in Appendix 8-13.

9.5.5 ESM assessment used in study 4

An important limitation of study 4 is the potential of measurement reactivity associated with asking participants to complete the same questionnaire multiple times. Although research has found reactivity to be fairly low in ESM studies (Cruise, Broderick, Porter, Kaell, & Stone, 1996; Ebner-Priemer & Sawitzki, 2007) it cannot be ruled out. Persistently asking patients to answer questions about their thoughts and feelings may lead to an increased focus of attention on their thoughts and feelings. According to the S-REF model, this is one of the exact processes believed to cause and maintain engagement in RNT and subsequent distress. Although the potential of measurement reactivity is inherent with all ESM studies, it is still important to consider when reflecting on the findings.

A final issue in study 4 relates to the questions used in the ESM assessment. Assessing the validity and reliability of ESM assessments is difficult (Palmier-Claus et al., 2011). The aim of ESM is to accurately capture recent or current experiences that may be have been missed or falsely reported when using traditional retrospective questionnaires. Thus, it makes little sense to attempt to validate ESM assessments by comparing ESM responses to those in traditional retrospective questionnaires. Another aim of ESM is to capture experiences as they fluctuate over time. Thus, it also makes little sense to assess the reliability (i.e. stability over time) of an ESM assessment. To

account for these difficulties, questions included in the ESM assessment in study 4 were informed by previous ESM assessments, feedback regarding the questions was obtained from a group of BCa survivors, and the questions were piloted in a group of BCa nurses.

9.6 Future research

While study 4 provides valuable first evidence to support the utility of the S-REF model for predicting RNT in BCa survivors, it did not provide support for the utility of the S-REF model in predicting distress. To determine whether this is because the S-REF model is insufficient for understanding distress in BCa survivors, or instead due to the low levels of distress amongst the sample, replication in a more distressed sample is needed. Moreover, as the smartphone ESM study focused only on BCa survivors who had completed treatment, the study needs to be replicated in a more heterogeneous sample to establish the generalisability of findings and the feasibility of smartphone ESM in BCa patients at other points in the disease trajectory.

In study 3, around 10% of treated and control patients in the total sample (i.e. distressed and non-distressed patients) deteriorated. While analyses were conducted on the sub-sample of patients who were distressed pre-treatment, analyses were not conducted on the sub-sample who were not distressed pre-treatment. However, around half of patients included in study 3 were not distressed (which was unsurprising as study 2 found that only 11% of RCTs screen for distress). From an ethical perspective, an evaluation of the proportion of non-distressed patients deteriorating is needed to ensure that providing psychological treatment to BCa patients with little distress (common amongst RCTs in this population) does not cause more harm than good.

More generally, future research should follow the Medical Research Council (MRC) framework for the development and evaluation of complex interventions (Craig et al., 2008), considered best practice for intervention development. The MRC framework proposes a systematic approach to the development of interventions that are supported by the best available evidence and appropriate theory. The framework includes four phases (Figure 9.2):

Phase 1: Development

The first phase of the MRC framework involves identifying what is already known about other interventions in the field to ensure new interventions are warranted. If no recent high-quality systematic review of the relevant literature exists, one should be conducted. Following this, it is important to explore the literature to ensure a theoretical basis exists for the proposed intervention. The clinical utility of the proposed theoretical model should then be tested and modelled through empirical studies.

Phase 2

Once the utility of the theoretical model guiding the intervention has been confirmed, the feasibility of the intervention and the proposed methods to evaluate it should then be tested. This includes testing procedures for their acceptability, estimating the likely rates of participant recruitment and retention, testing the quality of intervention delivery, and identifying any adverse events.

Phase 3

Phase three involves evaluating the efficacy of the proposed intervention. This ideally involves conducting an RCT. Conducting a process evaluation and/or an economic evaluation is also advised to help assess therapist adherence and competence to the intervention and provide important information regarding its cost-effectiveness.

Phase 4

The final phase involves translating the intervention into routine practice. This should be informed by the work carried out in all previous phases. Findings from the previous phases, both positive and negative, should be disseminated. A detailed description of the intervention should also be made available to allow replication. This phase also requires long-term follow-up of previous trials to determine whether short-term changes persist.

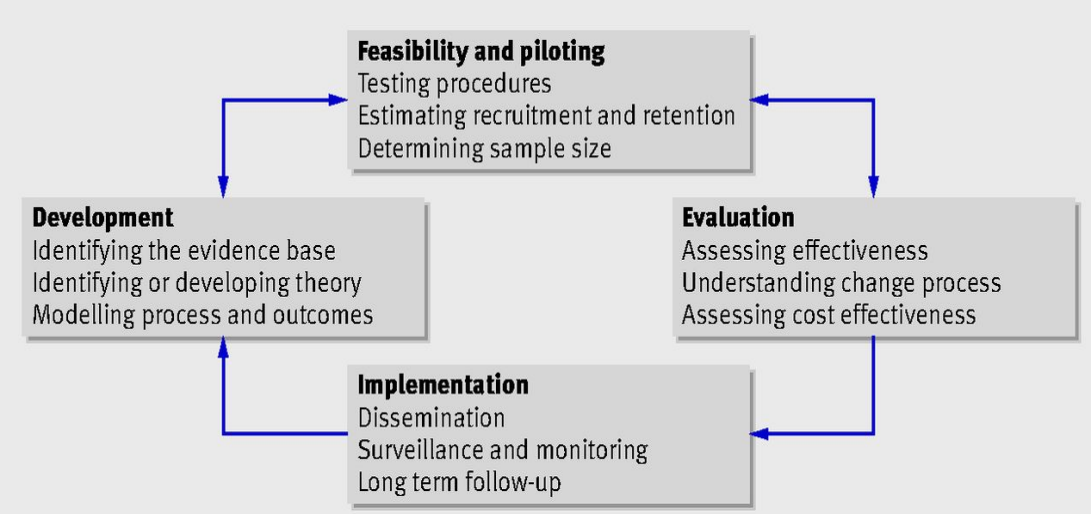


Figure 9.1: Key elements of the MRC development and evaluation process

The work in this thesis fulfils most of the aims in phase one - two high quality systematic reviews and an IPD-MA were conducted to identify whether efficacious psychological treatments for emotional distress in BCa exist; the literature was then searched to identify the evidence-base supporting the mechanisms of change proposed by the S-REF model and the IU model; and a smartphone ESM study was conducted to test the utility of the S-REF model and the IU model for predicting RNT and emotional distress in BCa survivors.

Future research should therefore continue to test the utility of the S-REF model and IU model for emotional distress in BCa or, after exploring the literature to ensure a theoretical basis for an alternative model exists, test the utility of an alternative model through a series of empirical studies. Once the psychological processes involved in the development and maintenance of emotional distress in BCa have been identified and findings have been replicated (as causality is never derived from a single study but through a variety of studies using different methods), therapeutic approaches targeting the identified processes should be tested by following phases 2-4 of the MRC framework.

References

Addis, M. E., Cardemil, E. V., Duncan, B. L., & Miller, S. D. (2006). Does manualization improve therapy outcomes? *Evidence-based practices in mental health: Debate and dialogue on the fundamental questions.* (pp. 131-160). Washington, DC, US: American Psychological Association.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control, 19*(6), 716-723.

Aldao, A. (2013). The future of emotion regulation research: Capturing context. *Perspectives on Psychological Science, 8*(2), 155-172.

Andrykowski, M. A., & Kangas, M. (2010). Posttraumatic stress disorder associated with cancer diagnosis and treatment. *Psycho-Oncology*, 348-357.

Ankuta, G. Y., & Abeles, N. (1993). Client satisfaction, clinical significance, and meaningful change in psychotherapy. *Professional Psychology: Research and Practice, 24*(1), 70-74.

Armes, J., Crowe, M., Colbourne, L., Morgan, H., Murrells, T., Oakley, C., . . . Richardson, A. (2009). Patients' supportive care needs beyond the end of cancer treatment: a prospective, longitudinal survey. *Journal of Clinical Oncology, 27*(36), 6172-6179.

Armijo-Olivo, S., Da Costa, B. R., Cummings, G. G., Ha, C., Fuentes, J., Saltaji, H., & Egger, M. (2015). PEDro or Cochrane to assess the quality of clinical trials? A meta-epidemiological study. *PloS One, 10*(7), e0132634.

Armijo-Olivo, S., Fuentes, J., Ospina, M., Saltaji, H., & Hartling, L. (2013). Inconsistency in the items included in tools used in general health research and physical therapy to evaluate the methodological quality of randomized controlled trials: a descriptive analysis. *BMC Medical Research Methodology, 13*(1), 116.

Arnberg, A., & Öst, L.-G. (2014). CBT for children with depressive symptoms: a meta-analysis. *Cognitive Behaviour Therapy, 43*(4), 275-288.

Atkins, D. C., Bedics, J. D., McGlinchey, J. B., & Beauchaine, T. P. (2005). Assessing clinical significance: Does it matter which method we use? *Journal of Consulting and Clinical Psychology, 73*(5), 982-989.

Baker, P., Beesley, H., Fletcher, I., Ablett, J., Holcombe, C., & Salmon, P. (2016). 'Getting back to normal'or 'a new type of normal'? A qualitative study of patients' responses to the existential threat of cancer. *European Journal of Cancer Care, 25*(1), 180-189.

Band, R., Barrowclough, C., Caldwell, K., Emsley, R., & Wearden, A. (2017). Activity patterns in response to symptoms in patients being treated for chronic fatigue syndrome: An experience sampling methodology study. *Health Psychology, 36*(3), 264-269.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin, 133*(1), 1-24.

Barber, J. P., Sharpless, B. A., Klostermann, S., & McCarthy, K. S. (2007). Assessing intervention competence and its relation to therapy outcome: A selected review derived from the outcome literature. *Professional Psychology: Research and Practice, 38*(5), 493-500.

Barlow, D., & Mavissakalian, M. (1981). Directions in the assessment and treatment of phobia: The next decade. *Phobia: Psychological and Pharmacological Treatment*, 199-245.

Barth, J., Munder, T., Gerger, H., Nüesch, E., Trelle, S., Znoj, H., . . . Cuijpers, P. (2016). Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *Focus, 14*(2), 229-243.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Bauer, S., Lambert, M. J., & Nielsen, S. L. (2004). Clinical significance methods: A comparison of statistical techniques. *Journal of Personality Assessment, 82*(1), 60-70.

Beatty, L., Kemp, E., Butow, P., Girgis, A., Schofield, P., Turner, J., . . . Koczwara, B. (2018). A systematic review of psychotherapeutic interventions for women with metastatic breast cancer: Context matters. *Psycho-Oncology, 27*(1), 34-42. doi: 10.1002/pon.4445

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. Pennsylvania: University of Pennsylvania Press.

Beck, A. T. (1976). *Cognitive Therapy and the Emotional Disorders*. New York: International Universities Press.

Behar, E., DiMarco, I. D., Hekler, E. B., Mohlman, J., & Staples, A. M. (2009). Current theoretical models of generalized anxiety disorder (GAD): Conceptual review and treatment implications. *Journal of Anxiety Disorders, 23*(8), 1011-1023.

Bennett-Levy, J., Butler, G., Fennell, M., Meuller, M., Westbrook, D. (2004). *Oxford Guide to Behavioural Experiments in Cognitive Therapy*. Oxford: Oxford University Press

Berkel, N. V., Ferreira, D., & Kostakos, V. (2017). The experience sampling method on mobile devices. *ACM Computing Surveys (CSUR), 50*(6), 1-40. doi: 10.1145/3123988

Berlin, J. A., & Rennie, D. (1999). Measuring the quality of trials: the quality of quality scales. *JAMA, 282*(11), 1083-1085.

Biau, D. J., Kernéis, S., & Porcher, R. (2008). Statistics in brief: the importance of sample size in the planning and interpretation of medical research. *Clinical Orthopaedics and Related Research, 466*(9), 2282-2288.

Bohlmeijer, E., Prenger, R., Taal, E., & Cuijpers, P. (2010). The effects of mindfulness-based stress reduction therapy on mental health of adults with a chronic medical disease: a meta-analysis. *Journal of Psychosomatic Research, 68*(6), 539-544.

Bolier, L., Haverman, M., Westerhof, G. J., Riper, H., Smit, F., & Bohlmeijer, E. (2013). Positive psychology interventions: a meta-analysis of randomized controlled studies. *BMC Public Health, 13*(1), 119.

Borkovec, T., Ray, W. J., & Stober, J. (1998). Worry: A cognitive phenomenon intimately linked to affective, physiological, and interpersonal behavioral processes. *Cognitive Therapy and Research, 22*(6), 561-576.

Borkovec, T., & Roemer, L. (1995). Perceived functions of worry among generalized anxiety disorder subjects: Distraction from more emotionally distressing topics. *Journal of Behavior Therapy and Experimental Psychiatry, 26*(1), 25-30.

Borkovec, T. D. (1994). *The Nature, Functions, and Origins of Worry. In: G. Davey & F. Tallis (Eds.)*. Sussex, England: Wiley & Sons.

Borkovec, T. D., Alcaine, O. M., & Behar, E. (2004). Avoidance theory of worry and generalized anxiety disorder in R. G. Heimberg, C. L. Turk, & D. S. Mennin (Eds.), *Generalized Anxiety Disorder: Advances in Research and Practice.* (pp. 77-108). New York, NY, US: Guilford Press.

Borkovec, T. D., Robinson, E., Pruzinsky, T., & DePree, J. A. (1983). Preliminary exploration of worry: Some characteristics and processes. *Behaviour Research and Therapy, 21*(1), 9-16.

Bouchard, L. C., Antoni, M. H., Blomberg, B. B., Stagl, J. M., Gudenkauf, L. M., Jutagir, D. R., . . . Derhagopian, R. P. (2016). Postsurgical depressive symptoms and proinflammatory cytokine elevations in women undergoing primary treatment for breast cancer. *Psychosomatic Medicine, 78*(1), 26-37.

Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., & Ravaud, P. (2008). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine, 148*(4), 295-309.

Boyes, A. W., Girgis, A., D'Este, C. A., Zucca, A. C., Lecathelinais, C., & Carey, M. L. (2013). Prevalence and predictors of the short-term trajectory of anxiety and depression in the first year after a cancer diagnosis: A population-based longitudinal study. *Journal of Clinical Oncology, 31*(21), 2724-2729.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 68*(6), 394-424. doi: 10.3322/caac.21492

Brennan, J., & Moynihan, C. (2004). *Cancer in Context: A Practical Guide to Supportive Care*. Oxford: Oxford University Press.

Brett, J., Staniszewska, S., Mockford, C., Herron-Marx, S., Hughes, J., Tysall, C., & Suleman, R. (2014). Mapping the impact of patient and public involvement on health and social care research: a systematic review. *Health Expectations, 17*(5), 637-650.

Broen, M. P., Marsman, V. A., Kuijf, M. L., Van Oostenbrugge, R. J., van Os, J., & Leentjens, A. F. (2016). Unraveling the relationship between motor symptoms, affective states and contextual factors in parkinson's disease: a feasibility study of the experience sampling method. *PloS One, 11*(3), e0151195.

Buhr, K., & Dugas, M. J. (2002). The intolerance of uncertainty scale: Psychometric properties of the English version. *Behaviour Research and Therapy, 40*(8), 931-945.

Buhr, K., & Dugas, M. J. (2009). The role of fear of anxiety and intolerance of uncertainty in worry: An experimental manipulation. *Behaviour Research and Therapy, 47*(3), 215-223.

Bultz, B. D., & Carlson, L. E. (2006). Emotional distress: the sixth vital sign-future directions in cancer care. *Psycho-Oncology, 15*(2), 93-95.

Burgess, C., Cornelius, V., Love, S., Graham, J., Richards, M., & Ramirez, A. (2005). Depression and anxiety in women with early breast cancer: five year observational cohort study. *Bmj, 330*(7493), 702.

Burgin, C. J., Silvia, P. J., Eddington, K. M., & Kwapil, T. R. (2013). Palm or cell? Comparing personal digital assistants and cell phones for experience sampling research. *Social Science Computer Review, 31*(2), 244-251.

Bylsma, L. M., Taylor-Clift, A., & Rottenberg, J. (2011). Emotional reactivity to daily events in major and minor depression. *Journal of Abnormal Psychology, 120*(1), 155-167.

Calmes, C. A., & Roberts, J. E. (2007). Repetitive thought and emotional distress: Rumination and worry as prospective predictors of depressive and anxious symptomatology. *Cognitive Therapy and Research, 31*(3), 343-356.

Cancer Research UK. (2015). Retrieved 07, 2018, from https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#heading-Eight

Carleton, R. N. (2012). The intolerance of uncertainty construct in the context of anxiety disorders: Theoretical and practical perspectives. *Expert Review of Neurotherapeutics, 12*(8), 937-947.

Carlson, L. E., & Bultz, B. D. (2004). Efficacy and medical cost offset of psychosocial interventions in cancer care: making the case for economic analyses. *Psycho-Oncology, 13*(12), 837-849.

Cerin, E., Szabo, A., & Williams, C. (2001). Is the experience sampling method (ESM) appropriate for studying pre-competitive emotions? *Psychology of Sport and Exercise, 2*(1), 27-45.

Chambless, D. L., & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology, 66*(1), 7-18.

Chan, M. W., Ho, S. M., Tedeschi, R. G., & Leung, C. W. (2011). The valence of attentional bias and cancer-related rumination in posttraumatic stress and posttraumatic growth among women with breast cancer. *Psycho-Oncology, 20*(5), 544-552.

Charache, H. (1932). Occupational and psychotherapy in the treatment of cancer. *Medical Journal & Record*.

Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behav Ther, 17*, 305-308.

Clark, H. D., Wells, G. A., Huët, C., McAlister, F. A., Salmi, L. R., Fergusson, D., & Laupacis, A. (1999). Assessing the quality of randomized trials: reliability of the Jadad scale. *Controlled Clinical Trials, 20*(5), 448-452.

Cobeanu, O., & David, D. (2018). Alleviation of side effects and distress in breast cancer patients by cognitive-behavioral interventions: A systematic review and meta-analysis. *Journal of Clinical Psychology in Medical Settings*, 1-21.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. 2nd. Hillsdale, NJ: Erlbaum.

Coleman, N., Hession, N., & Connolly, A. (2011). Psycho-oncology best practice guidelines and a service perspective: Conceptualising the fit and towards bridging the gap. *The Irish Journal of Psychology, 32*(1-2), 72-89. doi: 10.1080/03033910.2011.613980

Colleoni, M., Mandala, M., Peruzzotti, G., Robertson, C., Bredart, A., & Goldhirsch, A. (2000). Depression and degree of acceptance of adjuvant cytotoxic drugs. *The Lancet, 356*(9238), 1326-1327.

Cook, S. A., Salmon, P., Dunn, G., & Fisher, P. (2014). Measuring metacognition in cancer: validation of the metacognitions questionnaire 30 (MCQ-30). *PloS One, 9*(9), e107302.

Cook, S. A., Salmon, P., Dunn, G., Holcombe, C., Cornford, P., & Fisher, P. (2015a). The association of metacognitive beliefs with emotional distress after diagnosis of cancer. *Health Psychology, 34*(3), 207.

Cook, S. A., Salmon, P., Dunn, G., Holcombe, C., Cornford, P., & Fisher, P. (2015b). A prospective study of the association of metacognitive beliefs and processes with persistent emotional distress after diagnosis of cancer. *Cognitive Therapy and Research, 39*(1), 51-60.

Cook, S. A., Salmon, P., Hayes, G., Byrne, A., & Fisher, P. L. (2018). Predictors of emotional distress a year or more after diagnosis of cancer: A systematic review of the literature. *Psycho-Oncology, 27*(3), 791-801.

Cook., T. D., & Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis for Field Settings*. Boston, MA: Houghton-Mifflin.

Cooper, H., & Patall, E. A. (2009). The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychological Methods, 14*(2), 165-176.

Costa-Requena, G., Rodríguez, A., Fernández, R., Palomera, E., & Gil, F. L. (2011). Cognitive processing variables in breast cancer: Worry and distress at the end of treatment. *Journal of Cancer Education, 26*(2), 375-379.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ, 337*, a1655.

Crits-Christoph, P., Baranackie, K., Kurcias, J., Beck, A., Carroll, K., Perry, K., . . . Thompson, L. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research, 1*(2), 81-91.

Cruise, C. E., Broderick, J., Porter, L., Kaell, A., & Stone, A. A. (1996). Reactive effects of diary self-assessment in chronic pain patients. *Pain, 67*(2-3), 253-258.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease, 175*(9), 526-536. doi: 10.1097/00005053-198709000-00004

Csikszentmihalyi, M., & Larson, R. (2014). Validity and reliability of the experience-sampling method. *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* (pp. 35-54). Dordrecht: Springer Netherlands.

Cuijpers, P., van Straten, A., Bohlmeijer, E., Hollon, S., & Andersson, G. (2010). The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. *Psychological Medicine, 40*(2), 211-223.

Curran, S. L., Beacham, A. O., & Andrykowski, M. A. (2004). Ecological momentary assessment of fatigue following breast cancer treatment. *Journal of Behavioral Medicine, 27*(5), 425-444.

D'Hudson, G., & Saling, L. L. (2010). Worry and rumination in older adults: Differentiating the processes. *Aging & Mental Health, 14*(5), 524-534.

Dauchy, S., Léger, I., des Guetz, G., Ellien, F., Tidjani, L., Zelek, L., . . . Spano, J.-P. (2012). Quelle prise en charge psychologique pour les patients âgés atteints de cancer? *Psycho-Oncologie, 6*(1), 43-49.

Dawson, L., Zarin, D. A., Emanuel, E. J., Friedman, L. M., Chaudhari, B., & Goodman, S. N. (2009). Considering usual medical care in clinical trial design. *PLoS Medicine, 6*(9), e1000111.

Dekker, J., Braamse, A., Schuurhuizen, C., Beekman, A. T., van Linde, M., Sprangers, M. A., & Verheul, H. M. (2017). Distress in patients with cancer–on the need to distinguish between adaptive and maladaptive emotional responses. *Acta Oncologica, 56*(7), 1026-1029.

Delespaul, P. (1995). *Assessing Schizophrenia in Daily Life: The Experience Sampling Method*: UPM, Universitaire Pers Maastricht.

Delle Fave, A., & Massimini, F. (2004). Parenthood and the quality of experience in daily life: A longitudinal study. *Social Indicators Research, 67*(1-2), 75-106.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*(3), 177-188.

Deshields, T., Tibbs, T., Fan, M. Y., & Taylor, M. (2006). Differences in patterns of depression after treatment for breast cancer. *Psycho-Oncology, 15*(5), 398-406.

Diggle, P., Diggle, P. J., Heagerty, P., Heagerty, P. J., Liang, K.-Y., & Zeger, S. (2002). *Analysis of Longitudinal Data*. New York, NY: Oxford University Press.

Dugas, M. J., Buhr, K., & Ladouceur, R. (2004). The role of intolerance of uncertainty in etiology and maintenance *Generalized Anxiety Disorder: Advances in Research and Practice.* (pp. 143-163). New York, NY, US: Guilford Press.

Dugas, M. J., Gagnon, F., Ladouceur, R., & Freeston, M. H. (1998). Generalized anxiety disorder: A preliminary test of a conceptual model. *Behaviour Research and Therapy, 36*(2), 215-226.

Duijts, S. F., Faber, M. M., Oldenburg, H. S., van Beurden, M., & Aaronson, N. K. (2011). Effectiveness of behavioral techniques and physical exercise on psychosocial functioning and health-related quality of life in breast cancer patients and survivors: a meta-analysis. *Psycho-Oncology, 20*(2), 115-126.

Dupuy, J.-B., Beaudoin, S., Rhéaume, J., Ladouceur, R., & Dugas, M. (2001). Worry: Daily self-report in clinical and non-clinical populations. *Behaviour Research and Therapy, 39*(10), 1249-1255.

Ebner-Priemer, U. W., & Sawitzki, G. (2007). Ambulatory assessment of affective instability in borderline personality disorder. *European Journal of Psychological Assessment, 23*(4), 238-247.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629-634.

Eisenberg, S. A., Kurita, K., Taylor-Ford, M., Agus, D. B., Gross, M. E., & Meyerowitz, B. E. (2015). Intolerance of uncertainty, cognitive complaints, and cancer-related distress in prostate cancer survivors. *Psycho-Oncology, 24*(2), 228-235.

Exler, A., Kramer, S., Martínez, M. A. M., Navolskyi, C., Vogt, M., & Beigl, M. (2017). Suitability of event-based prompts in experience sampling studies focusing on location changes. *Paper Presented at the International Conference on Wireless Mobile Communication and Healthcare (pp. 163-168). Springer, Cham.*

Fahrenberg, J., Myrtek, M., Pawlik, K., & Perrez, M. (2007). Ambulatory assessment-monitoring behavior in daily life settings. *European Journal of Psychological Assessment, 23*(4), 206-213.

Farrington, D. P. (2003). Methodological quality standards for evaluation research. *The Annals of the American Academy of Political and Social Science, 587*(1), 49-68.

Fergus, T. A., & Wheless, N. (2018). Examining incremental explanatory power in accounting for worry severity: negative metacognitive beliefs uniquely predict worry severity following a worry episode. *Anxiety, Stress, & Coping*, 1-12.

Fisher, P. L., McNicol, K., Cherry, M. G., Young, B., Smith, E., Abbey, G., & Salmon, P. (2018). The association of metacognitive beliefs with emotional distress and trauma symptoms in adolescent and young adult survivors of cancer. *Journal of Psychosocial Oncology*, 1-12.

Fors, Bertheussen, G. F., Thune, I., Juvet, L. K., Elvsaas, I. K. Ø., Oldervoll, L., . . . Leivseth, G. (2011). Psychosocial interventions as part of breast cancer rehabilitation programs? Results from a systematic review. *Psycho-Oncology, 20*(9), 909-918.

Frans, Ö., Rimmö, P. A., Åberg, L., & Fredrikson, M. (2005). Trauma exposure and post-traumatic stress disorder in the general population. *Acta Psychiatrica Scandinavica, 111*(4), 291-290.

Freedland, K. E., Mohr, D. C., Davidson, K. W., & Schwartz, J. E. (2011). Usual and unusual care: existing practice control groups in randomized controlled trials of behavioral interventions. *Psychosomatic Medicine, 73*(4), 323-335.

Freeston, M. H., Rhéaume, J., Letarte, H., Dugas, M. J., & Ladouceur, R. (1994). Why do people worry? *Personality and Individual Differences, 17*(6), 791-802.

Fresco, D. M., Frankel, A. N., Mennin, D. S., Turk, C. L., & Heimberg, R. G. (2002). Distinct and overlapping features of rumination and worry: The relationship of cognitive production to negative affective states. *Cognitive Therapy and Research, 26*(2), 179-188.

Fritzsche, A., Dahme, B., Gotlib, I., Joormann, J., Magnussen, H., Watz, H., . . . von Leupoldt, A. (2010). Specificity of cognitive biases in patients with current depression and remitted depression and in patients with asthma. *Psychological Medicine, 40*(5), 815-826.

Frühauf, S., Gerger, H., Schmidt, H. M., Munder, T., & Barth, J. (2013). Efficacy of psychological interventions for sexual dysfunction: a systematic review and meta-analysis. *Archives of Sexual Behavior, 42*(6), 915-933.

Garfield, F. B., & Garfield, J. M. (2000). Clinical judgment and clinical practice guidelines. *International Journal of Technology Assessment in Health Care, 16*(04), 1050-1060.

Gellatly, J., Bower, P., Hennessy, S., Richards, D., Gilbody, S., & 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.

Gerber, A. J., Kocsis, J. H., Milrod, B. L., Roose, S. P., Barber, J. P., Thase, M. E., . . . Leon, A. C. (2011). A quality-based review of randomized controlled trials of psychodynamic psychotherapy. *American Journal of Psychiatry, 168*(1), 19-28. doi: 10.1176/appi.ajp.2010.08060843

Gerlach, A. L., Andor, T., & Patzelt, J. (2008). Die bedeutung von unsicherheitsintoleranz für die generalisierte angststörung modellüberlegungen und entwicklung einer deutschen version der unsicherheitsintoleranz-skala. *Zeitschrift für Klinische Psychologie und Psychotherapie, 37*(3), 190-199.

Gibbons, A., Groarke, A., & Sweeney, K. (2016). Predicting general and cancer-related distress in women with newly diagnosed breast cancer. *BMC Cancer, 16*(1), 935.

Girgis, A., Smith, A. B., & Durcinoska, I. (2018). Screening for distress in survivorship. *Current Opinion in Supportive and Palliative Care, 12*(1), 86-91.

Goldberg, D. (1992). *General health questionnaire -12 (GHQ-12)*. Windsor, England: NFER-Nelson

Goodman, S. N., & Berlin, J. A. (1994). The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. *Annals of Internal Medicine, 121*(3), 200-206.

Gorin, A. A., & Stone, A. A. (2001). Recall biases and cognitive errors in retrospective self-reports: A call for momentary assessments. *Handbook of Health Psychology, 23*, 405-413.

Goring, H. J., & Papageorgiou, C. (2008). Rumination and worry: Factor analysis of self-report measures in depressed participants. *Cognitive Therapy and Research, 32*(4), 554-566.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment, 26*(1), 41-54.

Groarke, A., Curtis, R., & Kerin, M. (2013). Cognitive-behavioural stress management enhances adjustment in women with breast cancer. *British Journal of Health Psychology, 18*(3), 623-641.

Hageman, W., & Arrindell, W. A. (1999). Establishing clinically significant change: Increment of precision and the distinction between individual and group level of analysis. *Behaviour Research and Therapy, 37*(12), 1169-1193.

Haller, H., Winkler, M. M., Klose, P., Dobos, G., Kummel, S., & Cramer, H. (2017). Mindfulness-based interventions for women with breast cancer: an updated systematic review and meta-analysis. *Acta Oncologica, 56*(12), 1665-1676. doi: 10.1080/0284186x.2017.1342862

Hann, D., Winter, K., & Jacobsen, P. (1999). Measurement of depressive symptoms in cancer patients: evaluation of the Center for Epidemiological Studies Depression Scale (CES-D). *Journal of Psychosomatic Research, 46*(5), 437-443.

Hare, D. J., Gracey, C., & Wood, C. (2016). Anxiety in high-functioning autism: A pilot study of experience sampling using a mobile platform. *Autism, 20*(6), 730-743.

Hartley, S., Haddock, G., Vasconcelos e Sa, D., Emsley, R., & Barrowclough, C. (2014). An experience sampling study of worry and rumination in psychosis. *Psychological Medicine, 44*(8), 1605-1614.

Hartling, L., Ospina, M., Liang, Y., Dryden, D. M., Hooton, N., Seida, J. K., & Klassen, T. P. (2009). Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. *BMJ, 339*, b4012.

Hashim, Z. (2016). Reliability and Validatidity of Hospital Anxiety and Depression Scale (HADS) on breast cancer survivors: Malaysia case study. Asia Pacific Environmental and Occupational Health Journal, 2(1).

Hassan, E. (2006). Recall bias can be a threat to retrospective and prospective research designs. *The Internet Journal of Epidemiology, 3*(2), 339-412.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience Sampling Method: Measuring the Quality of Everyday Life*. Thousand Oaks: Sage.

Helgeson, V. S., Snyder, P., & Seltman, H. (2004). Psychological and physical adjustment to breast cancer over 4 years: identifying distinct trajectories of change. *Health Psychology, 23*(1), 3-15.

Henselmans, I., Helgeson, V. S., Seltman, H., de Vries, J., Sanderman, R., & Ranchor, A. V. (2010). Identification and prediction of distress trajectories in the first year after a breast cancer diagnosis. *Health Psychology, 29*(2), 160-168.

Hewitt, M., Herdman, R., & Holland, D. (2004). *Meeting Psychosocial Needs of Women with Breast Cancer*. Wasington, DC: National Academies Press.

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., . . . Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ, 343*, d5928.

Higgins, J. P., & Green, S. (2011, March 2011). Cochrane handbook for systematic reviews of interventions Version 5.1.0. *The Cochrane Collaboration.* from www.cochrane-handbook.org

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal, 327*(7414), 557.

Holland, J. (1999). NCCN practice guidelines for the management of psychosocial distress. *Oncology, 13*(5A), 113-147.

Holland, J., Watson, M., & Dunn, J. (2011). The IPOS new International Standard of Quality Cancer Care: integrating the psychosocial domain into routine care. *Psycho-Oncology, 20*(7), 677-680.

Hollon, S. D., & Flick, S. N. (1988). On the meaning and methods of clinical significance. *Behavioral Assessment, 2*(10).

Holmes, W., Stewart, P., Garrow, A., Anderson, I., & Thorpe, L. (2002). Researching Aboriginal health: experience from a study of urban young people's health and well-being. *Social Science & Medicine, 54*(8), 1267-1279.

Horner, M., Ries, L., Krapcho, M., Neyman, N., Aminou, R., Howlader, N., . . . Mariotto, A. (2009). SEER Cancer Statistics Review, 1975-2006, National Cancer Institute. Bethesda, MD.

Howell, D., Currie, S., Mayo, S., Jones, G., Boyle, M., Hack, T., . . . Collacutt, V. (2009). A Pan-Canadian clinical practice guideline: Assessment of psychosocial health care needs of the adult cancer patient. *Toronto: Canadian Partnership Against Cancer (Cancer Journey Action Group) and the Canadian Association of Psychosocial Oncology*.

Howlader, N., Noone, A., & Krapcho, M. (2015). SEER Cancer Statistics Review, 1975-2012 [seer. cancer. gov/csr. 1975_2012/, based on the November 2014 SEER data submission, posted to the SEER website April 2015]. *Bethesda, MD: National Cancer Institute*.

Hsu, L. M. (1989). Reliable changes in psychotherapy: Taking into account regression toward the mean. *Behavioral Assessment, 4*(11).

Hsu, L. M. (1996). On the identification of clinically significant client changes: Reinterpretation of Jacobson's cut scores. *Journal of Psychopathology and Behavioral Assessment, 18*(4), 371-385.

Hufford, M. R., Shiffman, S., Paty, J., & Stone, A. A. (2001). Ecological momentary assessment: Real-world, real-time measurement of patient experience *Progress in Ambulatory Assessment: Computer-Assisted Psychological and Psychophysiological Methods in Monitoring and Field Studies.* (pp. 69-92). Ashland, OH, US: Hogrefe & Huber Publishers.

Huhn, M., Tardy, M., Spineli, L. M., Kissling, W., Förstl, H., Pitschel-Walz, G., . . . Davis, J. M. (2014). Efficacy of pharmacotherapy and psychotherapy for adult

psychiatric disorders: a systematic overview of meta-analyses. *JAMA Psychiatry, 71*(6), 706-715.

Hundt, N. E., Barrera, T. L., Robinson, A., & Cully, J. A. (2014). A systematic review of cognitive behavioral therapy for depression in Veterans. *Military Medicine, 179*(9), 942-949.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15*(4), 336-352.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1986). Toward a standard definition of clinically significant change. *Behavior Therapy, 17*(3), 308-311.

Jacobson, N. S., & Revenstorf, D. (1988). Statistics for assessing the clinical significance of psychotherapy techniques: Issues, problems, and new developments. *Behavioral Assessment, 2*(10), 133-145.

Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: description, application, and alternatives. *Journal of Consulting and Clinical Psychology, 67*(3), 300-307.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12-19.

Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J. M., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials, 17*(1), 1-12.

Jansson, L., & Öst, L.-G. (1982). Behavioral treatments for agoraphobia: An evaluative review. *Clinical Psychology Review, 2*(3), 311-336.

Jassim, G. A., Whitford, D. L., Hickey, A., & Carter, B. (2015). Psychological interventions for women with non-metastatic breast cancer. *Cochrane Database of Systematic Reviews*. doi: 10.1002/14651858.CD008729.pub2.

Jauhar, S., McKenna, P., Radua, J., Fung, E., Salvador, R., & Laws, K. (2014). Cognitive-behavioural therapy for the symptoms of schizophrenia: systematic review and meta-analysis with examination of potential bias. *The British Journal of Psychiatry, 204*(1), 20-29.

Johnson, E. I., Grondin, O., Barrault, M., Faytout, M., Helbig, S., Husky, M., . . . Wittchen, H. U. (2009). Computerized ambulatory monitoring in psychiatry: a multi-site collaborative study of acceptability, compliance, and reactivity. *International Journal of Methods in Psychiatric Research, 18*(1), 48-57.

Johnston, M., Pollard, B., & Hennessey, P. (2000). Construct validation of the hospital anxiety and depression scale with clinical populations. *Journal of Psychosomatic Research, 48*(6), 579-584.

Jüni, P., Witschi, A., Bloch, R., & Egger, M. (1999). The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA, 282*(11), 1054-1060.

Kang, M., Ragan, B. G., & Park, J.-H. (2008). Issues in outcomes research: an overview of randomization techniques for clinical trials. *Journal of Athletic Training, 43*(2), 215-221.

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification, 1*(4), 427-452.

Kazdin, A. E., & Wilson, G. T. (1978). *Evaluation of Behavior therapy: Issues, Evidence, and Research Strategies*: Cambridge, Massachusetts: Ballinger.

Khawaja, N. G., & McMahon, J. (2011). The relationship of meta-worry and intolerance of uncertainty with pathological worry, anxiety, and depression. *Behaviour Change, 28*(4), 165-180.

Kim, D.-M., Wampold, B. E., & Bolt, D. M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research, 16*(02), 161-172.

Kimhy, D., Delespaul, P., Corcoran, C., Ahn, H., Yale, S., & Malaspina, D. (2006). Computerized experience sampling method (ESMc): assessing feasibility and validity among individuals with schizophrenia. *Journal of Psychiatric Research, 40*(3), 221-230.

Klein, E., Stone, W. N., Hicks, M. W., & Pritchard, I. L. (2003). Understanding dropouts. *Journal of Mental Health Counseling, 25*(2), 89-100.

Klein, J., Jacobs, R. H., & Reinecke, M. A. (2007). Cognitive-behavioral therapy for adolescent depression: a meta-analytic investigation of changes in effect-size estimates. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(11), 1403-1413.

Kocsis, J. H., Gerber, A. J., Milrod, B., Roose, S. P., Barber, J., Thase, M. E., . . . Leon, A. C. (2010). A new scale for assessing the quality of randomized clinical trials of psychotherapy. *Comprehensive Psychiatry, 51*(3), 319-324.

Koerner, N., & Dugas, M. J. (2006). A cognitive model of generalized anxiety disorder: The role of intolerance of uncertainty. *Worry and its psychological disorders: Theory, Assessment and Treatment*, 201-216.

Kurita, K., Garon, E. B., Stanton, A. L., & Meyerowitz, B. E. (2013). Uncertainty and psychological adjustment in patients with lung cancer. *Psycho-Oncology, 22*(6), 1396-1401.

Lachance, S., Ladouceur, R., & Dugas, M. (1999). Elements explaining the tendency to worry. *Applied Psychology-An International Review-Psychologie Appliquee-Revue Internationale, 48*(2), 187-196.

Ladouceur, R., Gosselin, P., & Dugas, M. J. (2000). Experimental manipulation of intolerance of uncertainty: A study of a theoretical model of worry. *Behaviour Research and Therapy, 38*(9), 933-941.

Lambert, M. J., & Ogles, B. M. (2009). Using clinical significance in psychotherapy outcome research: The need for a common procedure and validity data. *Psychotherapy Research, 19*(4-5), 493-501.

Latronico, N., Botteri, M., Minelli, C., Zanotti, C., Bertolini, G., & Candiani, A. (2002). Quality of reporting of randomised controlled trials in the intensive care literature. *Intensive Care Medicine, 28*(9), 1316-1323.

Laugesen, N., Dugas, M. J., & Bukowski, W. M. (2003). Understanding adolescent worry: The application of a cognitive model. *Journal of Abnormal Child Psychology, 31*(1), 55-64.

Lehto, R. H., & Cimprich, B. (2009). Worry and the formation of cognitive representations of illness in individuals undergoing surgery for suspected lung cancer. *Cancer Nursing, 32*(1), 2-10.

Lepage, M., Sergerie, K., Pelletier, M., & Harvey, P.-O. (2007). Episodic memory bias and the symptoms of schizophrenia. *The Canadian Journal of Psychiatry, 52*(11), 702-709.

Levine, M., & Ensom, M. H. (2001). Post hoc power analysis: an idea whose time has passed? *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 21*(4), 405-409.

Li, M., Kennedy, E. B., Byrne, N., Gérin-Lajoie, C., Katz, M. R., Keshavarz, H., . . . Green, E. (2016). Management of depression in patients with cancer: a clinical practice guideline. *Journal of Oncology Practice, 12*(8), 747-756.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., . . . Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine, 6*(7), e1000100.

Liebherz, S., Schmidt, N., & Rabung, S. (2016). How to assess the quality of psychotherapy outcome studies: A systematic review of quality assessment criteria. *Psychotherapy Research, 26*(5), 573-589.

Litz, B. T., Gray, M. J., Bryant, R. A., & Adler, A. B. (2002). Early intervention for trauma: Current status and future directions. *Clinical Psychology: Science and Practice, 9*(2), 112-134.

Lo-Fo-Wong, D. N., de Haes, H. C., Aaronson, N. K., van Abbema, D. L., den Boer, M. D., van Hezewijk, M., . . . Reyners, A. K. (2016). Predictors of enduring clinical distress in women with breast cancer. *Breast Cancer Research and Treatment, 158*(3), 563-572.

Loerinc, A. G., Meuret, A. E., Twohig, M. P., Rosenfield, D., Bluett, E. J., & Craske, M. G. (2015). Response rates for CBT for anxiety disorders: Need for standardized criteria. *Clinical Psychology Review, 42*, 72-82.

Lohr, K. N. (2002). Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of Life Research, 11*(3), 193-205.

Luborsky, L., & DeRubeis, R. J. (1984). The use of psychotherapy treatment manuals: A small revolution in psychotherapy research style. *Clinical Psychology Review, 4*(1), 5-14.

Lunnen, K. M., & Ogles, B. M. (1998). A multiperspective, multivariable evaluation of reliable change. *Journal of Consulting and Clinical Psychology, 66*(2), 400-410.

Macmillan Cancer Support. (2016). Macmillan Cancer Support. *The emotional impact of cancer* Retrieved 23/05/2018, 2018, from www.macmillan.org.uk/documents/support_material/getinvolved/campaigns/impact_of_cancer_english.pdf

Maddams, J., Brewster, D., Gavin, A., Steward, J., Elliott, J., Utley, M., & Møller, H. (2009). Cancer prevalence in the United Kingdom: estimates for 2008. *British Journal of Cancer, 101*(3), 541.

Mahoney, A. E., & McEvoy, P. M. (2012). A transdiagnostic examination of intolerance of uncertainty across anxiety and depressive disorders. *Cognitive Behaviour Therapy, 41*(3), 212-222.

Mahoney, A. E., McEvoy, P. M., & Moulds, M. L. (2012). Psychometric properties of the Repetitive Thinking Questionnaire in a clinical sample. *Journal of Anxiety Disorders, 26*(2), 359-367.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22*(4), 719-748.

Marco, C. A., & Suls, J. (1993). Daily stress and the trajectory of mood: Spillover, response assimilation, contrast, and chronic negative affectivity. *Journal of Personality and Social Psychology, 64*(6), 1053-1063.

Martinovich, Z., Saunders, S., & Howard, K. (1996). Some comments on "Assessing clinical significance". *Psychotherapy Research, 6*(2), 124-132.

McEvoy, P. M., & Mahoney, A. E. (2012). To be sure, to be sure: Intolerance of uncertainty mediates symptoms of various anxiety disorders and depression. *Behavior Therapy, 43*(3), 533-545.

McEvoy, P. M., Mahoney, A. E., & Moulds, M. L. (2010). Are worry, rumination, and post-event processing one and the same?: Development of the Repetitive Thinking Questionnaire. *Journal of Anxiety Disorders, 24*(5), 509-519.

McEvoy, P. M., Salmon, K., Hyett, M. P., Jose, P. E., Gutenbrunner, C., Bryson, K., & Dewhirst, M. (2017). Repetitive negative thinking as a transdiagnostic

predictor of depression and anxiety symptoms in adolescents. *Assessment*, 1073191117693923.

McGlinchey, J. B., Atkins, D. C., & Jacobson, N. S. (2002). Clinical significance methods: Which one to use and how useful are they? *Behavior Therapy, 33*(4), 529-550.

McGlinchey, J. B., & Jacobson, N. S. (1999). Clinically significant but impractical? A A response to Hageman and Arrindell. Behaviour Research and Therapy, 37, 1211–1217.

McNair, D. M. (1971). *Manual profile of mood states*: Educational & Industrial testing service. San Diego, CA: Educational and Industrial Testing Services.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the penn state worry questionnaire. *Behaviour Research and Therapy, 28*(6), 487-495.

Millar, K., Purushotham, A. D., McLatchie, E., George, W. D., & Murray, G. D. (2005). A 1-year prospective study of individual variation in distress, and illness perceptions, after treatment for breast cancer. *Journal of Psychosomatic Research, 58*(4), 335-342.

Miller, K. D., Siegel, R. L., Lin, C. C., Mariotto, A. B., Kramer, J. L., Rowland, J. H., . . . Jemal, A. (2016). Cancer treatment and survivorship statistics, 2016. *CA: A Cancer Journal for Clinicians, 66*(4), 271-289.

Min, Y. H., Lee, J. W., Shin, Y.-W., Jo, M.-W., Sohn, G., Lee, J.-H., . . . Ko, B. S. (2014). Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: a feasibility study. *Journal of Medical Internet Research, 16*(5):e135.

Misono, S., Weiss, N. S., Fann, J. R., Redman, M., & Yueh, B. (2008). Incidence of suicide in persons with cancer. *Journal of Clinical Oncology, 26*(29), 4731-4738.

Moberly, N. J., & Watkins, E. R. (2008). Ruminative self-focus and negative affect: an experience sampling study. *Journal of Abnormal Psychology, 117*(2), 314-323.

Modi, A. C., Wagner, J., Smith, A. W., Kellermann, T. S., & Michaelis, R. (2017). Implementation of psychological clinical trials in epilepsy: Review and guide. *Epilepsy & Behavior, 74*, 104-113.

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy, 36*(9), 809-848.

Mohr, D. C., Spring, B., Freedland, K. E., Beckner, V., Arean, P., Hollon, S. D., . . . Kaplan, R. (2009). The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and Psychosomatics, 78*(5), 275-284.

Mosavel, M., Simon, C., Van Stade, D., & Buchbinder, M. (2005). Community-based participatory research (CBPR) in South Africa: engaging multiple constituents to shape the research question. *Social Science & Medicine, 61*(12), 2577-2587.

Mujagic, Z., Leue, C., Vork, L., Lousberg, R., Jonkers, D., Keszthelyi, D., . . . Masclee, A. (2015). The Experience Sampling Method - a new digital tool for momentary symptom assessment in IBS: an exploratory study. *Neurogastroenterology & Motility, 27*(9), 1295-1302.

Munder, T., & Barth, J. (2018). Cochrane's risk of bias tool in the context of psychotherapy outcome research. *Psychotherapy Research, 28*(3), 347-355.

Muris, P., Roelofs, J., Rassin, E., Franken, I., & Mayer, B. (2005). Mediating effects of rumination and worry on the links between neuroticism, anxiety and depression. *Personality and Individual Differences, 39*(6), 1105-1111.

Muris, P., & van der Heiden, S. (2006). Anxiety, depression, and judgments about the probability of future negative and positive events in children. *Journal of Anxiety Disorders, 20*(2), 252-261.

Mustafa, M., Carson-Stevens, A., Gillespie, D., & Edwards, A. G. (2013). Psychological interventions for women with metastatic breast cancer. *Cochrane Database of Systematic Reviews*(6), CD004253. doi: 10.1002/14651858.CD004253.pub4

Myin-Germeys, I., Krabbendam, L., Jolles, J., Delespaul, P. A., & van Os, J. (2002). Are cognitive impairments associated with sensitivity to stress in schizophrenia? An experience sampling study. *American Journal of Psychiatry, 159*(3), 443-449.

Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & Van Os, J. (2009). Experience sampling research in psychopathology: opening the black box of daily life. *Psychological Medicine, 39*(9), 1533-1547.

Myin-Germeys, I., van Os, J., Schwartz, J. E., Stone, A. A., & Delespaul, P. A. (2001). Emotional reactivity to daily life stress in psychosis. *Archives of General Psychiatry, 58*(12), 1137-1144.

Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., & Reininghaus, U. (2018). Experience sampling methodology in mental health research: new insights and technical developments. *World Psychiatry, 17*(2), 123-132.

Naaman, S. C., Radwan, K., Fergusson, D., & Johnson, S. (2009). Status of psychological trials in breast cancer patients: a report of three meta-analyses. *Psychiatry: Interpersonal and Biological Processes, 72*(1), 50-69.

National Breast Cancer Centre. (2003). *Clinical Practice Guidelines for the Psychosocial Care of Adults with Cancer*: National Breast Cancer Centre, National Cancer Control Initiative.

National Comprehensive Cancer Network. (2003). Distress management. Clinical practice guidelines. *Journal of the National Comprehensive Cancer Network: JNCCN, 1*(3), 344-374.

National Institute for Clinical Excellence. (2004). *Guidance on Cancer Services: Improving Supportive and Palliative Care for Adults with Cancer: the manual*: National Institute for Clinical Excellence.

Newnham, E. A., Harwood, K. E., & Page, A. C. (2007). Evaluating the clinical significance of responses by psychiatric inpatients to the mental health subscales of the SF-36. *Journal of Affective Disorders, 98*(1-2), 91-97.

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology, 100*(4), 569-582.

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology, 109*(3), 504-511.

Norström, F. (2015). Poor quality in the reporting and use of statistical methods in public health–the case of unemployment and health. *Archives of Public Health, 73*(1), 56.

Nunnally, J. C., & Kotsch, W. E. (1983). Studies of individual subjects: Logic and methods of analysis. *British Journal of Clinical Psychology, 22*(2), 83-93.

Nyenhuis, D. L., Yamamoto, C., Luchetta, T., Terrien, A., & Parmentier, A. (1999). Adult and geriatric normative data and validation of the profile of mood states. *Journal of Clinical Psychology, 55*(1), 79-86.

Ofcom. (2018). UK: Smartphone ownership by age from 2012-2018 *Statista - The Statistics Portal.* 2018, from https://www.statista.com/statistics/271851/smartphone-owners-in-the-united-kingdom-uk-by-age/.

Ogles, B. M., Lambert, M. J., & Sawyer, J. D. (1995). Clinical significance of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Journal of Consulting and Clinical Psychology, 63*(2), 321-326.

Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical Psychology Review, 21*(3), 421-446.

Okediji, P. T., Salako, O., & Fatiregun, O. O. (2017). Pattern and predictors of unmet supportive care needs in cancer patients. *Cureus, 9*(5).e1234.

Olivo, S. A., Macedo, L. G., Gadotti, I. C., Fuentes, J., Stanton, T., & Magee, D. J. (2008). Scales to assess the quality of randomized controlled trials: a systematic review. *Physical Therapy, 88*(2), 156-175.

Öst, L.-G. (2008). Efficacy of the third wave of behavioral therapies: A systematic review and meta-analysis. *Behaviour Research and Therapy, 46*(3), 296-321.

Öst, L.-G. (2014). The efficacy of acceptance and commitment therapy: an updated systematic review and meta-analysis. *Behaviour Research and Therapy, 61*, 105-121.

Öst, L.-G., Havnen, A., Hansen, B., & Kvale, G. (2015). Cognitive behavioral treatments of obsessive–compulsive disorder. A systematic review and meta-analysis of studies published 1993–2014. *Clinical Psychology Review, 40*, 156-169.

Öst, L.-G., & Ollendick, T. H. (2017). Brief, intensive and concentrated cognitive behavioral treatments for anxiety disorders in children: A systematic review and meta-analysis. *Behaviour Research and Therapy, 97*, 134-145.

Öst, L.-G., Riise, E. N., Wergeland, G. J., Hansen, B., & Kvale, G. (2016). Cognitive behavioral and pharmacological treatments of OCD in children: A systematic review and meta-analysis. *Journal of Anxiety Disorders, 43*, 58-69.

Page, A. E., & Adler, N. E. (2008). *Cancer Care for the Whole Patient: Meeting Psychosocial Health Needs*. Washington, DC: National Academies Press.

Palmier-Claus, J. E., Ainsworth, J., Machin, M., Barrowclough, C., Dunn, G., Barkus, E., . . . Buchan, I. (2012). The feasibility and validity of ambulatory self-report of psychotic symptoms using a smartphone software application. *BMC Psychiatry, 12*(1), 172.

Palmier-Claus, J. E., Myin-Germeys, I., Barkus, E., Bentley, L., Udachina, A., Delespaul, P., . . . Dunn, G. (2011). Experience sampling research in individuals with mental illness: reflections and guidance. *Acta Psychiatrica Scandinavica, 123*(1), 12-20.

Palmier-Claus, J. E., Taylor, P. J., Ainsworth, J., Machin, M., Dunn, G., & Lewis, S. W. (2014). The temporal association between self-Injurious thoughts and psychotic symptoms: A mobile phone assessment study. *Suicide and Life-Threatening Behavior, 44*(1), 101-110.

Papageorgiou, C., & Wells, A. (Eds.). (2004). *Depressive Rumination: Nature, Rheory, and Rreatment.* New York: Wiley

Parsons, H. M. (1974). What Happened at Hawthorne?: New evidence suggests the Hawthorne effect resulted from operant reinforcement contingencies. *Science,* 183(4128), 922-932.

Peeters, F., Nicolson, N. A., Berkhof, J., Delespaul, P., & deVries, M. (2003). Effects of daily events on mood states in major depressive disorder. *Journal of Abnormal Psychology, 112*(2), 203-211.

Perepletchikova, F., Treat, T. A., & Kazdin, A. E. (2007). Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. *Journal of Consulting and Clinical Psychology, 75*(6), 829-841.

Quattropani, M. C., Lenzo, V., Mucciardi, M., & Toffle, M. E. (2015). The role of metacognitions in predicting anxiety and depression levels in cancer patients ongoing chemotherapy. *Procedia-Social and Behavioral Sciences, 205*, 463-473.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401.

Reddy, M. (2010). Suicide incidence and epidemiology. *Indian Journal of Psychological Medicine, 32*(2), 77-82.

Reese, C., Weis, J., Schmucker, D., & Mittag, O. (2017). Development of practice guidelines for psychological interventions in the rehabilitation of patients with oncological disease (breast, prostate, or colorectal cancer): Methods and results. *Psycho-Oncology, 26*(10), 1513-1518.

Risley, T. R. (1970). Behavior modification: An experimental-therapeutic endeavor. *Behavior Modification and Ideal Mental Health Services*, 103-127.

Rissanen, R., Nordin, K., Ahlgren, J., & Arving, C. (2015). A stepped care stress management intervention on cancer-related traumatic stress symptoms among breast cancer patients—a randomized study in group vs. individual setting. *Psycho-Oncology, 24*(9), 1028-1035.

Robichaud, M., Dugas, M. J., & Conway, M. (2003). Gender differences in worry and associated cognitive-behavioral variables. *Journal of Anxiety Disorders, 17*(5), 501-516.

Ronk, F. R., Hooke, G. R., & Page, A. C. (2012). How consistent are clinical significance classifications when calculation methods and outcome measures differ? *Clinical Psychology: Science and Practice, 19*(2), 167-179.

Ronk, F. R., Hooke, G. R., & Page, A. C. (2016). Validity of clinically significant change classifications yielded by Jacobson-Truax and Hageman-Arrindell methods. *BMC Psychiatry, 16*(1), 187.

Rood, L., Roelofs, J., Bögels, S. M., & Alloy, L. B. (2010). Dimensions of negative thinking and the relations with symptoms of depression and anxiety in children and adolescents. *Cognitive Therapy and Research, 34*(4), 333-342.

Rose, S., Bisson, J., & Wessely, S. (2003). A systematic review of single-session psychological interventions ('debriefing') following trauma. *Psychotherapy and Psychosomatics, 72*(4), 176-184.

Rosen, N. O., & Knäuper, B. (2009). A little uncertainty goes a long way: State and trait differences in uncertainty interact to increase information seeking but also increase worry. *Health Communication, 24*(3), 228-238.

Rothwell, P. M. (2005). External validity of randomised controlled trials:"to whom do the results of this trial apply?". *The Lancet, 365*(9453), 82-93.

Rozental, A., Magnusson, K., Boettcher, J., Andersson, G., & Carlbring, P. (2017). For better or worse: An individual patient data meta-analysis of deterioration

among participants receiving Internet-based cognitive behavior therapy. *Journal of Consulting and Clinical Psychology, 85*(2), 160-177.

Ruscio, A. M., Gentes, E. L., Jones, J. D., Hallion, L. S., Coleman, E. S., & Swendsen, J. (2015). Rumination predicts heightened responding to stressful life events in major depressive disorder and generalized anxiety disorder. *Journal of Abnormal Psychology, 124*(1), 17-26.

Salander, P. (2017). Does advocating screening for distress in cancer rest more on ideology than on science? *Patient Education and Counseling, 100*(5), 858-860.

Salmon, P., Clark, L., McGrath, E., & Fisher, P. (2015). Screening for psychological distress in cancer: renewing the research agenda. *Psycho-Oncology, 24*(3), 262-268.

Sanson-Fisher, R., Girgis, A., Boyes, A., Bonevski, B., Burton, L., Cook, P., & Group, S. C. R. (2000). The unmet supportive care needs of patients with cancer. *Cancer, 88*(1), 226-237.

Saunders, S. M., Howard, K. I., & Newman, F. L. (1988). Evaluating the clinical significance of treatment effects: Norms and normality. *Behavioral Assessment*, 10(2), 207-218.

Savard, J., Simard, S., Ivers, H., & Morin, C. M. (2005). Randomized study on the efficacy of cognitive-behavioral therapy for insomnia secondary to breast cancer, part I: Sleep and psychological effects. *Journal of Clinical Oncology, 23*(25), 6083-6096.

Schaeffeler, N., Pfeiffer, K., Ringwald, J., Brucker, S., Wallwiener, M., Zipfel, S., & Teufel, M. (2015). Assessing the need for psychooncological support: screening instruments in combination with patients' subjective evaluation may define psychooncological pathways. *Psycho-Oncology, 24*(12), 1784-1791.

Schellekens, M., van den Hurk, D., Prins, J., Donders, A., Molema, J., Dekhuijzen, R., . . . Speckens, A. (2017). Mindfulness-based stress reduction added to care as usual for lung cancer patients and/or their partners: A multicentre randomized controlled trial. *Psycho-Oncology, 26*(12), 2118-2126.

Schneider, B., Scissons, H., Arney, L., Benson, G., Derry, J., Lucas, K., . . . Sunderland, M. (2004). Communication between people with schizophrenia and their medical professionals: A participatory research project. *Qualitative Health Research, 14*(4), 562-577.

Segerstrom, S. C., Tsao, J. C., Alden, L. E., & Craske, M. G. (2000). Worry and rumination: Repetitive thought as a concomitant and predictor of negative mood. *Cognitive Therapy and Research, 24*(6), 671-688.

Seggar, L. B., Lambert, M. J., & Hansen, N. B. (2002). Assessing clinical significance: application to the Beck Depression Inventory. *Behavior Therapy, 33*(2), 253-269.

Seidel, M., Petermann, J., Diestel, S., Ritschel, F., Boehm, I., King, J. A., . . . Goschke, T. (2016). A naturalistic examination of negative affect and disorder-related rumination in anorexia nervosa. *European Child & Adolescent Psychiatry, 25*(11), 1207-1216.

Sheldrick, R. C., Kendall, P. C., & Heimberg, R. G. (2001). The clinical significance of treatments: A comparison of three treatments for conduct disordered children. *Clinical Psychology: Science and Practice, 8*(4), 418-430.

Siegel, R. L., Miller, K. D., & Jemal, A. (2017). Cancer statistics, 2017. *CA: A cancer Journal for Clinicians, 67*(1), 7-30.

Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. *CA: A Cancer Journal for Clinicians, 68*(1), 7-30. doi: doi:10.3322/caac.21442.

Simmonds, A. (2010). Ethics of placebo-controlled trials in developing countries: The search for standards and solutions. *The Journal of Undergraduate Writing Program, New York: Columbia University*, 1-7.

Sloan, E., Hall, K., Moulding, R., Bryce, S., Mildred, H., & Staiger, P. K. (2017). Emotion regulation as a transdiagnostic treatment construct across anxiety, depression, substance, eating and borderline personality disorders: a systematic review. *Clinical Psychology Review*.

Smelt, A. F., van der Weele, G. M., Blom, J. W., Gussekloo, J., & Assendelft, W. J. (2010). How usual is usual care in pragmatic intervention studies in primary care? An overview of recent trials. *British Journal of General Practice, 60*(576), e305-e318.

Soo, H., & Sherman, K. (2015). Rumination, psychological distress and post-traumatic growth in women diagnosed with breast cancer. *Psycho-Oncology, 24*(1), 70-79.

Souza, A. C. d., Alexandre, N. M. C., & Guirardello, E. d. B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Serviços de Saúde, 26*(3), 649-659.

Speer, D. C. (1992). Clinically significant change: Jacobson and Truax (1991) revisited. *Journal of Consulting and Clinical Psychology, 60*(3), 402-408.

Speer, D. C., & Greenbaum, P. E. (1995). Five methods for computing significant individual client change and improvement rates: Support for an individual growth curve approach. *Journal of Consulting and Clinical Psychology, 63*(6), 1044-1048.

Spinhoven, P., Drost, J., van Hemert, B., & Penninx, B. W. (2015). Common rather than unique aspects of repetitive negative thinking are related to depressive and anxiety disorders and symptoms. *Journal of Anxiety Disorders, 33*, 45-52.

Stanton, A. L., Ganz, P. A., Kwan, L., Meyerowitz, B. E., Bower, J. E., Krupnick, J. L., . . . Belin, T. R. (2005). Outcomes from the moving beyond cancer

psychoeducational, randomized, controlled trial with breast cancer patients. *Journal of Clinical Oncology, 23*(25), 6009-6018.

Stewart, L. A., & Parmar, M. K. (1993). Meta-analysis of the literature or of individual patient data: is there a difference? *The Lancet, 341*(8842), 418-422.

Stewart, L. A., & Tierney, J. F. (2002). To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Evaluation & the Health Professions, 25*(1), 76-97.

Stone, A. A., & Shiffman, S. (1994). Ecological momentary assessment (EMA) in behavorial medicine. *Annals of Behavioral Medicine*. 3(6), 199-202.

Stone, A. A., Shiffman, S., Schwartz, J. E., Broderick, J. E., & Hufford, M. R. (2002). Patient non-compliance with paper diaries. *BMJ, 324*(7347), 1193-1194.

Strupp, H. H., & Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcomes: With special reference to negative effects in psychotherapy. *American Psychologist, 32*(3), 187-196.

Stuart, A., & Ord, J. (1994). *JK Kendall's Advanced Theory of Statistics.* (6th ed.). London, England: Arnold Publishers.

Sullivan, T. P., Khondkaryan, E., Dos Santos, N. P., & Peters, E. N. (2011). Applying experience sampling methods to partner violence research: Safety and feasibility in a 90-day study of community women. *Violence against Women, 17*(2), 251-266.

Sundell, K., & Åhsberg, E. (2016). Trends in methodological quality in controlled trials of psychological and social interventions. *Research on Social Work Practice,* 28(5), 568–576.

Suresh, K. (2011). An overview of randomization techniques: an unbiased assessment of outcome in clinical research. *Journal of Human Reproductive Sciences, 4*(1), 8-11.

Swain, J., Hancock, K., Dixon, A., & Bowman, J. (2015). Acceptance and commitment therapy for children: A systematic review of intervention studies. *Journal of Contextual Behavioral Science, 4*(2), 73-85.

Swain, J., Hancock, K., Hainsworth, C., & Bowman, J. (2013). Acceptance and commitment therapy in the treatment of anxiety: a systematic review. *Clinical Psychology Review, 33*(8), 965-978.

Swift, J. K., Callahan, J., & Levine, J. C. (2009). Using clinically significant change to identify premature termination. *Psychotherapy: Theory, Research, Practice, Training, 46*(3), 328.

Taha, S. A., Matheson, K., & Anisman, H. (2012). Everyday experiences of women posttreatment after breast cancer: the role of uncertainty, hassles, uplifts, and coping on depressive symptoms. *Journal of Psychosocial Oncology, 30*(3), 359-379.

Tatrow, K., & Montgomery, G. H. (2006). Cognitive behavioral therapy techniques for distress and pain in breast cancer patients: a meta-analysis. *Journal of Behavioral Medicine, 29*(1), 17-27.

Temple, J., Salmon, P., Tudur-Smith, C., Huntley, C. D., & Fisher, P. L. (2018). A systematic review of the quality of randomized controlled trials of psychological treatments for emotional distress in breast cancer. *Journal of Psychosomatic Research, 108*, 22-31.

Thewes, B., Bell, M., & Butow, P. (2013). Fear of cancer recurrence in young early-stage breast cancer survivors: the role of metacognitive style and disease-related factors. *Psycho-Oncology, 22*(9), 2059-2063.

Thielsch, C., Andor, T., & Ehring, T. (2015a). Do metacognitions and intolerance of uncertainty predict worry in everyday life? An ecological momentary assessment study. *Behavior Therapy, 46*(4), 532-543.

Thielsch, C., Andor, T., & Ehring, T. (2015b). Metacognitions, intolerance of uncertainty and worry: An investigation in adolescents. *Personality and Individual Differences, 74*, 94-98.

Thoma, N. C., McKay, D., Gerber, A. J., Milrod, B. L., Edwards, A. R., & Kocsis, J. H. (2012). A quality-based review of randomized controlled trials of cognitive-behavioral therapy for depression: an assessment and metaregression. *American Journal of Psychiatry, 169*(1), 22-30.

Thomas, V., & Azmitia, M. (2016). Tapping into the app: Updating the experience sampling method for the 21st century. *Emerging Adulthood, 4*(1), 60-67.

Tierney, J. F., Vale, C., Riley, R., Smith, C. T., Stewart, L., Clarke, M., & Rovers, M. (2015). Individual participant data (IPD) meta-analyses of randomised controlled trials: guidance on their use. *PLoS Medicine, 12*(7), e1001855.

Tingey, R., Lambert, M., Burlingame, G., & Hansen, N. (1996a). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research, 6*(2), 109-123.

Tingey, R., Lambert, M., Burlingame, G., & Hansen, N. (1996b). Clinically significant change: Practical indicators for evaluating psychotherapy outcome. *Psychotherapy Research, 6*(2), 144-153.

Tit, A., Camilla, A., Angela, A., Marco, A., Gianni, A., Barceló, A. M., . . . Keith, C. (2017). *European Guide on Quality Improvement in Comprehensive Cancer Control*: National Institute of Public Health.

Tolin, D. (1999). A revised meta-analysis of psychosocial treatments for PTSD. *Poster Presented at AABT, Toronto*.

Tolin, D. F., McKay, D., Forman, E. M., Klonsky, E. D., & Thombs, B. D. (2015). Empirically supported treatment: Recommendations for a new model. *Clinical Psychology: Science and Practice, 22*(4), 317-338.

Tondorf, T., Grossert, A., Rothschild, S., Koller, M., Rochlitz, C., Kiss, A., . . . Zwahlen, D. (2018). Focusing on cancer patients' intentions to use psychooncological support: A longitudinal, mixed-methods study. *Psycho-oncology*, *27*(6), 1656-1663.

van der Velden, R. M., Mulders, A. E., Drukker, M., Kuijf, M. L., & Leentjens, A. F. (2018). Network analysis of symptoms in a Parkinson patient using experience sampling data: An n= 1 study. *Movement Disorders*. doi: 10.1002/mds.93

Verhagen, A. P., de Vet, H. C., de Bie, R. A., Boers, M., & van den Brandt, P. A. (2001). The art of quality assessment of RCTs included in systematic reviews. *Journal of Clinical Epidemiology, 54*(7), 651-654.

Villoria, E., & Lara, L. (2018). Assessment of a hospital anxiety and depression scale for cancer patients. Revista medica de Chile, 146(3).

Wampold, B. E., & Jenson, W. R. (1986). Clinical significance revisited. *Behavior Therapy, 17*(3), 302-305.

Weber, F., & Exner, C. (2013). Metacognitive beliefs and rumination: A longitudinal study. *Cognitive Therapy and Research, 37*(6), 1257-1261. doi: 10.1007/s10608-013-9555-y

Wells, A. (2000). *Emotional Disorders and Metacognition: Innovative Cognitive Therapy*. Chichester, UK: Wiley.

Wells, A. (2009). *Metacognitive Therapy for Anxiety and Depression*. New York: Guilford Press.

Wells, A., & Cartwright-Hatton, S. (2004). A short form of the metacognitions questionnaire: properties of the MCQ-30. *Behaviour Research and Therapy, 42*(4), 385-396.

Wells, A., & Fisher, P. (2015). *Treating Depression: MCT, CBT, and Third Wave Therapies*: Chichester, UK: Wiley. https://doi.org/10.1002/9781119114482

Wells, A., & Matthews, G. (1994). *Attention and Emotion. A clinical perspective*. Hove, UK: Erlbaum.

Wells, A., & Matthews, G. (1996). Modelling cognition in emotional disorder: The S-REF model. *Behaviour Research and Therapy, 34*(11-12), 881-888.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin, 130*(4), 631-663.

Westermann, S., Grezellschak, S., Oravecz, Z., Moritz, S., Lüdtke, T., & Jansen, A. (2017). Untangling the complex relationships between symptoms of schizophrenia and emotion dynamics in daily life: Findings from an experience sampling pilot study. *Psychiatry Research, 257*, 514-518.

Wijesekera, P., Baokar, A., Tsai, L., Reardon, J., Egelman, S., Wagner, D., & Beznosov, K. (2018). Dynamically regulating mobile application permissions. *IEEE Security & Privacy, 16*(1), 64-71.

Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*(2), 203-214.

Xiao, F., Song, X., Chen, Q., Dai, Y., Xu, R., Qiu, C., & Guo, Q. (2017). Effectiveness of psychological interventions on depression in patients after breast cancer surgery: a meta-analysis of randomized controlled trials. *Clinical Breast Cancer, 17*(3), 171-179.

Ye, M., Du, K., Zhou, J., Zhou, Q., Shou, M., Hu, B., . . . Liang, S. (2018). A meta-analysis of the efficacy of cognitive behavior therapy on quality of life and psychological health of breast cancer survivors and patients. *Psycho-Oncology*, 27(7), 1695-1703.

Zeng, X., Zhang, Y., Kwong, J. S., Zhang, C., Li, S., Sun, F., . . . Du, L. (2015). The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. *Journal of Evidence-Based Medicine, 8*(1), 2-10.

Zhang, J., Xu, R., Wang, B., & Wang, J. (2016). Effects of mindfulness-based therapy for patients with breast cancer: A systematic review and meta-analysis. *Complementary Therapies in Medicine, 26*, 1-10.

Zhang, Z. (2016). Multiple imputation with multivariate imputation by chained equation (MICE) package. *Annals of Translational Medicine, 4*(2), 30.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*(6), 361-370.

Zimmermann, T., Heinrichs, N., & Baucom, D. H. (2007). "Does one size fit all?" Moderators in psychosocial interventions for breast cancer patients: a meta-analysis. *Annals of Behavioral Medicine, 34*(3), 225-239.

List of Appendices

Appendix 1: Final search strategy used to search electronic databases (study 1)

Appendix 1A: Search strategy for PubMed

Number	Term
1	*Review* OR meta**
2	"Psychotherapy"[Mesh] OR "psychotherapy" OR "psychological therapy" OR "counselling" OR "counselling" OR "psychological intervention" OR "cognitive behavioural therapy" OR "group therapy" OR "psychosocial therapy" OR "psychological treatment" OR "individual therapy" OR "psychotherapeutic" OR "CBT"
3	"Breast Neoplasms"[mesh] OR "breast neoplasms" OR "breast cancer"
4	"Depression"[Mesh] OR "depressive disorder"[Mesh] OR "depressive disorder" OR "anx*"[Mesh] Or "anxiety disorders"[Mesh] OR "anxiety disorders" OR "anxiety" OR "depress*" OR "emotional distress" OR "psychological distress"
5	1 AND 2 AND 3 AND 4
6	Limit to English language

Abbreviations: MeSH, Medical Subject Headings; mh, Mesh Heading; Tiab, Title or Abstract; SH, Subject Heading

Appendix 1B: Search strategy for PsycINFO, Web of Science, Scopus, and the Cochrane Database of Systematic Reviews.

Number	Term
1	Review* OR meta*
2	"Psychotherapy" OR "psychological therapy" OR "counseling" OR "counselling" OR "psychological intervention" OR "cognitive behavioural therapy" OR "group therapy" OR "psychosocial therapy" OR "psychological treatment" OR "individual therapy" OR "psychotherapeutic" OR "CBT"
3	"Breast neoplasm*" OR "breast cancer"
4	"Depress*" OR "depressive disorder" OR "anx*" OR "anxiety disorder*" OR "emotional distress" OR "psychological distress"
5	1 AND 2 AND 3 AND 4
6	Limit to English language

Appendix 2: Final search strategy used to search electronic databases for study 2 & 3

Appendix 2A: Search strategy for PubMed

Number	Term
1	Randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR placebo[tiab] OR drug therapy[sh] OR randomly[tiab] OR trial[tiab] OR groups[tiab] NOT (animals [mh] NOT humans [mh]
2	"Psychotherapy"[Mesh] OR "psychotherapy" OR "psychological therapy" OR "counselling" OR "counselling" OR "psychological intervention" OR "cognitive behavioural therapy" OR "group therapy" OR "psychosocial therapy" OR "psychological treatment" OR "individual therapy" OR "psychotherapeutic" OR "CBT"
3	"Breast Neoplasms"[mesh] OR "breast neoplasms" OR "breast cancer"
4	"Depression"[Mesh] OR "depressive disorder"[Mesh] OR "depressive disorder" OR "anx*"[Mesh] Or "anxiety disorders"[Mesh] OR "anxiety disorders" OR "anxiety" OR "depress*" OR "emotional distress" OR "psychological distress"
5	"gene therapy" OR "genetic*"
6	1 AND 2 AND 3 AND 4 NOT 5
7	Limit to English language

Abbreviations: MeSH, Medical Subject Headings; mh, Mesh Heading; Tiab, Title or Abstract; SH, Subject Heading

Appendix 2B: Search strategy for PsycINFO, Web of Science, Scopus, PsycARTICLE, and AMED

Number	Term
1	"Psychotherapy" OR "psychological therapy" OR "counseling" OR "counselling" OR "psychological intervention" OR "cognitive behavioural therapy" OR "group therapy" OR "psychosocial therapy" OR "psychological treatment" OR "individual therapy" OR "psychotherapeutic" OR "CBT"
2	"Breast neoplasm*" OR "breast cancer"
3	"Depression" OR "depressive disorder" OR "anxiety" OR "anxiety disorder*" OR "emotional distress" OR "psychological distress"
4	"gene therapy" OR "genetic*"
5	1 AND 2 AND 3 NOT 4
6	Limit to English language

Abbreviations: MeSH, Medical Subject Headings; mh, Mesh Heading; Tiab, Title or Abstract; SH, Subject Heading

Appendix 3:

Standardised data extraction tool used to extract data from included studies (study 2 & 3)

Review Title:

Study authors:

Name of review author completing this form:

Date form completed:

STUDY METHODS

Aim of intervention (what was the problem that this intervention was designed to address?)	

POPULATION AND SETTING

Inclusion criteria	
Exclusion criteria	
Method/s of recruitment of participants (How were potential participants approached and invited to participate?)	

PARTICIPANTS

	Description as stated in report/paper
Total no. randomised (overall and per group)	
Number Invited to participate, eligible, excluded, refused to take part	
Number post randomisation (excluded, withdrawn, lost to follow-up)	
Age: range, mean and SD (Overall and per group, if available)	
Race/Ethnicity (Overall and per group, if available)	
Stage of disease **(Overall and per group, if available: Also if certain stages are part of inclusion criteria)**	
Treatment status (Overall and per group, if available)	
Did participants need to score above a particular distress threshold? If so, how was this justified?	

Psychological comorbidities: e.g. PTSD along with anxiety (Overall and per group, if available)	
Non-psychological comorbidities e.g. Other cancer along with BCa cancer (Overall and per group, if available)	

INTERVENTION 1

	Description as stated in report/paper
Intervention type (e.g. CBT, psychoeducation etc)	
Format (group/ individual)	
Number and length of sessions	
Delivery method (phone, internet, in person etc)	

INTERVENTION 2 (Or control) – copy and paste table for however many more groups

	Description as stated in report/paper
Control type (e.g. TAU, WLC, no treatment) If TAU – describe what this was	
Format (group/ individual)	
Number and length of sessions	
Delivery method (phone, internet, in person etc)	

OUTCOMES

	Description as stated in report/paper
Methods of assessing outcome measures (e.g., questionnaire or interview)	
Primary outcome(s)	
Secondary outcome (s)	
Distress outcome (e.g. anxiety, depression, mood)	
Distress measure	
27. Statistical significance between groups on distress outcome (if this is not primary outcome)? Yes – in favour of treatment No Yes- in favour of comparison	

Appendix 4:

Psychotherapy outcome study methodology rating form (POMRF)

Note: If not enough information is given regarding a specific item a rating of 0 is given.

1. Clarity of sample description

0	Poor	Vague description of sample (e.g. only mentioned whether patients were diagnosed with the disorder).
1	Fair	Fair description of sample (e.g. mentioned inclusion/exclusion criteria, demographics, etc.).
2	Good	Good description of sample (e.g. mentioned inclusion/exclusion criteria, demographics, and the prevalence of comorbid disorders).

2. Severity/chronicity of the disorder

0	Poor	Severity/chronicity was not reported and/or subsyndromal patients were included in the sample.
1	Fair	All patients met the criteria for the disorder. Sample includes acute (<1 yr) and/or low severity.
2	Good	Sample consisted entirely of chronic (>1 yr) patients of at least moderate severity.

3. Representativeness of sample

0	Poor	Sample is very different from patients seeking treatment for the disorder (e.g. there are excessively strict exclusion criteria).
1	Fair	Sample is somewhat representative of patients seeking treatment for the disorder (e.g. patients were only excluded if they met criteria for other major disorders).
2	Good	Sample is very representative of patients seeking treatment for the disorder (e.g. authors made efforts to ensure representativeness of sample).

4. Reliability of the diagnosis in question

0	Poor	The diagnostic process was not reported, or not assessed with structured interviews by a trained interviewer.
1	Fair	The diagnosis was assessed with structured interview by a trained interviewer.
2	Good	The diagnosis was assessed with structured interview by a trained interviewer and adequate inter-rater reliability was demonstrated (e.g. *kappa* coefficient).

5. Specificity of outcome measures		
0	Poor	Very broad outcome measures, not specific to the disorder (e.g. SCL-90R total score).
1	Fair	Moderately specific outcome measures
2	Good	Specific outcome measures, such as a measure for each symptom cluster.
6. Reliability and validity of outcome measures		
0	Poor	Measures have unknown psychometric properties, or properties that fail to meet current standards of acceptability.
1	Fair	Some, but not all measures have known or adequate psychometric properties.
2	Good	All measures have good psychometric properties. The outcome measures are the best available for the authors' purpose.
7. Use of blind evaluators		
0	Poor	Blind assessor was not used (e.g. assessor was the therapist, assessor was not blind to treatment condition, or the authors do not specify).
1	Fair	Blind assessor was used, but no checks were used to assess the blind.
2	Good	Blind assessor was used in correct fashion. Checks were used to assess whether the assessor was aware of treatment condition.
8. Assessor training		
0	Poor	Assessor training and accuracy are not specified, or are unacceptable.
1	Fair	Minimum criterion for assessor training is specified (e.g. assessor has had specific training in the use of the outcome measure), but accuracy is not monitored or reported.
2	Good	Minimum criterion of assessor training is specified. Inter-rater reliability was checked, and/or assessment procedures were calibrated during the study to prevent evaluator drift.
9. Assignment to treatment		
0	Poor	Biased assignment, e.g. patients selected their own therapy or were assigned in another non-random fashion, or there is only one group.

1	Fair	Random or stratified assignment. There may be some systematic bias but not enough to pose a serious threat to internal validity. There may be therapist by treatment confounds. *N* may be too small to protect against bias.
2	Good	Random or stratified assignment, and patients are randomly assigned to therapists within condition. When theoretically different treatments are used, each treatment is provided by a large enough number of different therapists. *N* is large enough to protect against bias.
10. Design		
0	Poor	Active treatment vs. WLC, or briefly described TAU.
1	Fair	Active treatment vs. TAU with good description, or placebo condition.
2	Good	Active treatment vs. another previously empirically documented active treatment.
11. Power analysis		
0	Poor	No power analysis was made prior to the initiation of the study.
1	Fair	A power analysis based on an estimated effect size was used.
2	Good	A data-informed power analysis was made and the sample size was decided accordingly.
12. Assessment points		
0	Poor	Only pre- and post-treatment, or pre- and follow-up.
1	Fair	Pre-, post-, and follow-up <1 year.
2	Good	Pre-, post-, and follow-up⩾1 year.
13. Manualised, replicable, specific treatment programs		
0	Poor	Description of treatment procedure is unclear, and treatment is not based on a publicly available, detailed treatment manual. Patients may be receiving multiple forms of treatment at once in an uncontrolled manner.
1	Fair	Treatment is not designed for the disorder, or description of the treatment is generally clear and based on a publicly available, detailed treatment manual, but there are some ambiguities about the procedure. Patients may have received additional forms of treatment, but this is balanced between groups or otherwise controlled.

2	Good	Treatment is designed for the disorder. A detailed treatment manual is available, and/or treatment is explained in sufficient detail for replication. No ambiguities about the treatment procedure. Patients receive only the treatment in question.
14. Number of therapists		
0	Poor	Only one therapist, i.e. complete confounding between therapy and therapist.
1	Fair	At least two therapists, but the effect of therapist on outcome is not analysed.
2	Good	Three, or more therapists, and the effect of therapist on outcome is analysed.
15. Therapist training/experience		
0	Poor	Very limited clinical experience of the treatment and/or disorder (e.g. students).
1	Fair	Some clinical experience of the treatment and/or disorder.
2	Good	Long clinical experience of the treatment and the disorder (e.g. practicing therapists).
16. Checks for therapist adherence		
0	Poor	No checks were made to assure that the intervention was consistent with protocol.
1	Fair	Some checks were made (e.g. assessed a proportion of therapy tapes).
2	Good	Frequent checks were made (e.g. weekly supervision of each session using a detailed rating form).
17. Checks for therapist competence		
0	Poor	No checks were made to assure that the intervention was delivered competently.
1	Fair	Some checks were made (e.g. assessed a proportion of therapy tapes).
2	Good	Frequent checks were made (e.g. weekly supervision of each session using a detailed rating form).
18. Control of concomitant treatments (e.g. medications)		

0	Poor	No attempt to control for concomitant treatments, or no information about concomitant treatments provided. Patients may have been receiving other forms of treatment in addition to the study treatment.
1	Fair	Asked patients to keep medications stable and/or to discontinue other psychological therapies during the treatment.
2	Good	Ensured that patients did not receive any other treatments (medical or psychological) during the study.
19. Handling of attrition		
0	Poor	Proportions of attrition are not described, or described but no dropout analysis is performed.
1	Fair	Proportions of attrition are described, and dropout analysis or intent-to-treat analysis is performed.
2	Good	No attrition, or proportions of attrition are described, dropout analysis is performed, and results are presented as intent-to-treat analysis.
20. Statistical analyses and presentation of results		
0	Poor	Inadequate statistical methods are used and/or data are not fully presented.
1	Fair	Adequate statistical methods are used but data are not fully presented.
2	Good	Adequate statistical methods are used and data are presented with *M* and SD.
21. Clinical significance		
0	Poor	No presentation of clinical significance was done.
1	Fair	An arbitrary criterion for clinical significance was used and the conditions were compared regarding percent clinically improved.
2	Good	Jacobson's criteria for clinical significance were used and presented for a selection (or all) of the outcome measures, and conditions were compared regarding percent clinically improved.
22. Equality of therapy hours (for non-WLC designs only)		
0	Poor	Conditions differ markedly (⩾20% difference in therapy hours).
1	Fair	Conditions differ somewhat (10–19% difference in therapy hours).
2	Good	Conditions do not differ (<10% difference in therapy hours).

Appendix 5:

References of the RCTs included (study 2)

References of Included RCTs

Abad, A. N. S., Bakhtiari, M., Kashani, F. L., & Habibi, M. (2016). The comparison of effectiveness of treatment based on acceptance and commitment with cognitive-behavioral therapy in reduction of stress and anxiety in cancer patients. International *Journal of Cancer Research and Prevention*, 9(3), 229-246.

Allard, N. C. (2007). Day surgery for breast cancer: effects of a psychoeducational telephone intervention on functional status and emotional distress. *Oncology Nursing Forum*, 34, 133–141

Allen, S. M., Shah, A. C., Nezu, A. M., Nezu, C. M., Ciambrone, D., Hogan, J., & Mor, V. (2002). A problem-solving approach to stress reduction among younger women with breast carcinoma. *Cancer*, 94(12), 3089-3100.

Andersen, B. L., Farrar, W. B., Golden-Kreutz, D. M., Glaser, R., Emery, C. F., Crespin, T. R., . . . Carson, W. E. (2004). Psychological, behavioral, and immune changes after a psychological intervention: A clinical trial. Journal of *Clinical Oncology*, 22(17), 3570-3580. doi:10.1200/jco.2004.06.030

Anton, S., Gugic, D., Katinic, K., & Topic, J. (2015). Efficasy of Different Psychiatric Treatment Methods of Liaison Psychiatrist in Treatment of Women with Breast Cancer. *Collegium Antropologicum*, 39(2), 377-383.

Antoni, M. H., Lechner, S., Diaz, A., Vargas, S., Holley, H., Phillips, K., . . . Blomberg, B. (2009). Cognitive behavioral stress management effects on psychosocial and physiological adaptation in women undergoing treatment for breast cancer. *Brain, Behavior, and Immunity*, 23(5), 580-591. doi:10.1016/j.bbi.2008.09.005

Antoni, M. H., Lehman, J. M., Kilbourn, K. M., Boyers, A. E., Culver, J. L., Alferi, S. M., . . . Harris, S. D. (2001). Cognitive-behavioral stress management intervention decreases the prevalence of depression and enhances benefit finding among women under treatment for early-stage breast cancer. *Health Psychology*, 20(1), 20-32.

Antoni, M. H., Wimberly, S. R., Lechner, S. C., Kazi, A., Sifre, T., Urcuyo, K. R., . . . Carver, C. S. (2006). Reduction of cancer-specific thought intrusions and anxiety symptoms with a stress management intervention among women undergoing treatment for breast cancer. *American Journal of Psychiatry*, 163(10), 1791-1797. doi:10.1176/appi.ajp.163.10.1791

Arving, C., Sjödén, P. O., Bergh, J., Hellbom, M., Johansson, B., Glimelius, B., & Brandberg, Y. (2007). Individual psychosocial support for breast cancer patients: A randomized study of nurse versus psychologist interventions and standard care. *Cancer Nursing*, 30(3), E10-E19. doi:10.1097/01.NCC.0000270709.64790.05

Ashing, K., & Rosales, M. (2014). A telephonic-based trial to reduce depressive symptoms among Latina breast cancer survivors. *Psycho-Oncology*, 23(5), 507-515. doi:10.1002/pon.3441

Badger, T., Segrin, C., Dorros, S. M., Meek, P., & Lopez, A. M. (2007). Depression and Anxiety in Women with Breast Cancer and their Partners. *Nursing Research*, 56(1), 44-53. doi:10.1097/00006199-200701000-00006

Badger, T., Segrin, C., Pasvogel, A., & Lopez, A. M. (2013a). The effect of psychosocial interventions delivered by telephone and videophone on quality of life in early-stage breast cancer survivors and their supportive partners. *Journal of Telemedicine and Telecare*, 19(5), 260-265. doi:10.1177/1357633x13492289

Badger, T. A., Segrin, C., Hepworth, J. T., Pasvogel, A., Weihs, K., & Lopez, A. M. (2013b). Telephone-delivered health education and interpersonal counseling improve quality of life for Latinas with breast cancer and their supportive partners. *Psycho-Oncology*, 22(5), 1035-1042. doi:10.1002/pon.3101

Baucom, D. H., Porter, L. S., Kirby, J. S., Gremore, T. M., Wiesenthal, N., Aldridge, W., . . . Keefe, F. J. (2009). A couple-based intervention for female breast cancer. *Psycho-Oncology*, 18(3), 276-283. doi:10.1002/pon.1395

Beutel, M. E., Weissflog, G., Leuteritz, K., Wiltink, J., Haselbacher, A., Ruckes, C., . . . Braehler, E. (2014). Efficacy of short-term psychodynamic psychotherapy (STPP) with depressed breast cancer patients: results of a randomized controlled multicenter trial. *Annals of Oncology*, 25(2), 378-384. doi:10.1093/annonc/mdt526

Björneklett, H. G., Lindemalm, C., Rosenblad, A., Ojutkangas, M. L., Letocha, H., Strang, P., & Bergkvist, L. (2012). A randomised controlled trial of support group intervention after breast cancer treatment: Results on anxiety and depression. *Acta Oncologica*, 51(2), 198-207. doi:10.3109/0284186X.2011.610352

Boesen, E. H., Karlsen, R., Christensen, J., Paaschburg, B., Nielsen, D., Bloch, I. S., . . . Johansen, C. (2011). Psychosocial group intervention for patients with primary breast cancer: A randomised trial. *European Journal of Cancer*, 47(9), 1363-1372. doi:10.1016/j.ejca.2011.01.002

Bower, J. E., Crosswell, A. D., Stanton, A. L., Crespi, C. M., Winston, D., Arevalo, J., . . . Ganz, P. A. (2015). Mindfulness meditation for younger breast cancer survivors: a randomized controlled trial. *Cancer*, 121(8), 1231-1240. doi:10.1002/cncr.29194

Bredal, I. S., Karesen, R., Smeby, N. A., Espe, R., Sorensen, E. M., Amundsen, M., . . . Ekeberg, O. (2014). Effects of a Psychoeducational Versus a Support Group Intervention in Patients With Early-Stage Breast Cancer Results of a Randomized Controlled Trial. *Cancer Nursing*, 37(3), 198-207. doi:10.1097/NCC.0b013e31829879a3

Budin, W. C., Hoskins, C. N., Haber, J., Sherman, D. W., Maislin, G., Cater, J. R., . . . Shukla, S. (2008). Breast cancer: Education, counseling, and adjustment among patients and partners: A randomized clinical trial. *Nursing Research,* 57(3), 199-213. doi:10.1097/01.nnr.0000319496.67369.37

Burton, M. V., Parker, R. W., Farrell, A., Bailey, D., Conneely, J., Booth, S., & Elcombe, S. (1995). A randomized controlled trial of preoperative psychological preparation for mastectomy. *Psycho-Oncology,* 4(1), 1-19. doi:10.1002/pon.2960040102

Carlson, L. E., Doll, R., Stephen, J., Faris, P., Tamagawa, R., Drysdale, E., & Speca, M. (2013). Randomized controlled trial of mindfulness-based cancer recovery versus supportive expressive group therapy for distressed survivors of breast cancer (MINDSET). *Journal of Clinical Oncology,* 31(25), 3119-3126.

Chan, C. L., Ho, R. T., Lee, P. W., Cheng, J. Y., Leung, P. P., Foo, W., . . . Spiegel, D. (2006). A randomized controlled trial of psychosocial interventions using the psychophysiological framework for Chinese breast cancer patients. *Journal of Psychosocial Oncology,* 24(1), 3-26.

Christensen, D. N. (1983). Post-mastectomy couple counseling - an outcome study of a structured treatment protocol. *Journal of Sex & Marital Therapy,* 9(4), 266-275.

Classen, C., Butler, L. D., Koopman, C., Miller, E., DiMiceli, S., Giese-Davis, J., . . . Spiegel, D. (2001). Supportive-expressive group therapy and distress in patients with metastatic breast cancer - A randomized clinical intervention trial. *Archives of General Psychiatry,* 58(5), 494-501. doi:10.1001/archpsyc.58.5.494

Classen, C. C., Kraemer, H. C., Blasey, C., Giese-Davis, J., Koopman, C., Palesh, O. G., . . . Spiegel, D. (2008). Supportive-expressive group therapy for primary breast cancer patients: a randomized prospective multicenter trial. *Psycho-Oncology,* 17(5), 438-447. doi:10.1002/pon.1280.

Cohen, M., & Fried, G. (2007). Comparing relaxation training and cognitive-behavioral group therapy for women with breast cancer. *Research on Social Work Practice,* 17(3), 313-323. doi:10.1177/1049731506293741

Cousson-Gelie, F., Bruchon-Schweitzer, M., Atzeni, T., & Houede, N. (2011). Evaluation of a psychosocial intervention on social support, perceived control, coping strategies, emotional distress and quality of life of breast cancer patients. *Psychological Reports,* 108(3), 923-942. doi:10.2466/02.07.15.20.PR0.108.3.923-942

Cruess, D. G., Antoni, M. H., McGregor, B. A., Kilbourn, K. M., Boyers, A. E., Alferi, S. M., . . . Kumar, M. (2000). Cognitive-behavioral stress management reduces serum cortisol by enhancing benefit finding among women being treated for early stage breast cancer. *Psychosomatic Medicine,* 62(3), 304-308.

Davis, H. I. (1986). Effects of biofeedback and cognitive therapy on stress in patients with breast cancer. *Psychological Reports,* 59(2), 967-974.

Dirksen, S. R., & Epstein, D. R. (2008). Efficacy of an insomnia intervention on fatigue, mood and quality of life in breast cancer survivors. *Journal of Advanced Nursing,* 61(6), 664-675. doi:10.1111/j.1365-2648.2007.04560.x

Dolbeault, S., Cayrou, S., Brédart, A., Viala, A. L., Desclaux, B., Saltel, P., . . . Dickes, P. (2009). The effectiveness of a psycho-educational group after early-stage breast cancer treatment: Results of a randomized French study. *Psycho-Oncology,* 18(6), 647-656. doi:10.1002/pon.1440

Dowlatabadi, M. M., Ahmadi, S. M., Sorbi, M. H., Beiki, O., Razavi, T. K., & Bidaki, R. (2016). The effectiveness of group positive psychotherapy on depression and happiness in breast cancer patients: A randomized controlled trial. *Electronic Physician,* 8(3), 2175-2180. doi:10.19082/2175

Edelman, S., Bell, D. R., & Kidman, A. D. (1999). Group CBT versus supportive therapy with patients who have primary breast cancer. *Journal of Cognitive Psychotherapy,* 13(3), 189-202.

Edelman, S., Bell, D. R., & Kidman, A. D. (1999). A group cognitive behaviour therapy programme with metastatic breast cancer patients. *Psycho-Oncology: Journal of the Psychological, Social and Behavioral Dimensions of Cancer*, *8*(4), 295-305.

Edmonds, C. V. I. (1998). A randomized study assessing the psychological effects of a long-term psychological intervention on metastatic breast cancer patients. (58), ProQuest Information & Learning, US. Retrieved from http://search.ebscohost.com.ezproxy.liv.ac.uk/login.aspx?direct=true&db=psyh&AN=1998-95002-004&site=ehost-live&scope=site Available from EBSCOhost psyh database.

Ferguson, R. J., McDonald, B. C., Rocque, M. A., Furstenberg, C. T., Horrigan, S., Ahles, T. A., & Saykin, A. J. (2012). Development of CBT for chemotherapy-related cognitive change: results of a waitlist control trial. *Psycho-Oncology*, *21*(2), 176-186.

Ferguson, R. J., Sigmon, S. T., Pritchard, A. J., Labrie, S. L., Goetze, R. E., Fink, C. M., & Garrett, A. M. (2016). A randomized trial of videoconference-delivered cognitive behavioral therapy for survivors of breast cancer with self-reported cognitive dysfunction. *Cancer,* 122(11), 1782-1791. doi:10.1002/cncr.29891

Fillion, L., Gagnon, P., Leblond, F., Gelinas, C., Savard, J., Dupuis, R., . . . Larochelle, M. (2008). A brief intervention for fatigue management in breast cancer survivors. *Cancer Nursing,* 31(2), 145-159. doi:10.1097/01.ncc.0000305698.97625.95

Fukui, S., Kugaya, A., Okamura, H., Kamiya, M., Koike, M., Nakanishi, T., ... & Uchitomi, Y. (2000). A psychosocial group intervention for Japanese women with primary breast carcinoma: a randomized controlled trial. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 89(5), 1026-1036.

Gaston-Johansson, F., Fall-Dickson, J. M., Nanda, J., Ohly, K. V., Stillman, S., Krumm, S., & Kennedy, M. J. (2000). The effectiveness of the comprehensive coping strategy program on clinical outcomes in breast cancer autologous bone marrow transplantation. *Cancer Nursing,* 23(4), 277-285.

Giese-Davis, J., Bliss-Isberg, C., Wittenberg, L., White, J., Star, P., Zhong, L., . . . Spiegel, D. (2016). Peer-counseling for women newly diagnosed with breast cancer: A randomized community/research collaboration trial. Cancer, 122(15), 2408-2417.

Goodwin, P. J., Leszcz, M., Ennis, M., Koopmans, J., Vincent, L., Guther, H., . . . Navarro, M. (2001). The effect of group psychosocial support on survival in metastatic breast cancer. *New England Journal of Medicine,* 345(24), 1719-1726.

Gotay, C. C., Moinpour, C. M., Unger, J. M., Jiang, C. S., Coleman, D., Martino, S., . . . Albain, K. S. (2007). Impact of a peer-delivered telephone intervention for women experiencing a breast cancer recurrence. *Journal of Clinical Oncology,* 25(15), 2093-2099. doi:10.1200/jco.2006.07.4674

Graves, K. D., Carter, C. L., Anderson, E. S., & WInett, R. A. (2003). Quality of life pilot intervention for breast cancer patients: use of social cognitive theory. *Palliative Support Care,* 1(02), 121-134.

Groarke, A., Curtis, R., & Kerin, M. (2013). Cognitive-behavioural stress management enhances adjustment in women with breast cancer. British *Journal of Health Psychology,* 18(3), 623-641. doi:10.1111/bjhp.12009

Gudenkauf, L. M., Antoni, M. H., Stagl, J. M., Lechner, S. C., Jutagir, D. R., Bouchard, L. C., . . . Giron, G. L. (2015). Brief cognitive–behavioral and relaxation training interventions for breast cancer: A randomized controlled trial. *Journal of Consulting and Clinical Psychology,* 83(4), 677.

Heiney, S. P., McWayne, J., Hurley, T. G., Lamb Jr, L. S., Bryant, L. H., Butler, W., & Godder, K. (2003). Efficacy of therapeutic group by telephone for women with breast cancer. *Cancer Nursing,* 26(6), 439-447.

Henderson, V. P., Clemow, L., Massion, A. O., Hurley, T. G., Druker, S., & Hebert, J. R. (2012). The effects of mindfulness-based stress reduction on psychosocial outcomes and quality of life in early-stage breast cancer patients: a randomized trial. *Breast Cancer Research and Treatment,* 131(1), 99-109. doi:10.1007/s10549-011-1738-1

Hoffman, C. J., Ersser, S. J., Hopkinson, J. B., Nicholls, P. G., Harrington, J. E., & Thomas, P. W. (2012). Effectiveness of mindfulness-based stress reduction in mood, breast- and endocrine-related quality of life, and well-being in stage 0 to III breast cancer: a randomized, controlled trial. *Journal of Clinical Oncology,* 30(12), 1335-1342. doi:10.1200/jco.2010.34.0331

Hopko, D. R., Armento, M. E. A., Robertson, S. M. C., Ryba, M. M., Carvalho, J. P., Colman, L. K., . . . Lejuez, C. W. (2011). Brief behavioral activation and problem-solving therapy for depressed breast cancer patients: Randomized

trial. *Journal of Consulting and Clinical Psychology*, 79(6), 834-849. doi:10.1037/a0025450

Kalaitzi, C., Papadopoulos, V. P., Michas, K., Vlasis, K., Skandalakis, P., & Filippou, D. (2007). Combined brief psychosexual intervention after mastectomy: Effects on sexuality, body image, and psychological well-being. *Journal of Surgical Oncology*, 96(3), 235-240. doi:10.1002/jso.20811

Kissane, D. W., Bloch, S., Smith, G. C., Miach, P., Clarke, D. M., Ikin, J., . . . McKenzie, D. (2003). Cognitive-existential group psychotherapy for women with primary breast cancer: A randomized controlled trial. *Psycho-Oncology*, 12(6), 532-546. doi:10.1002/pon.683

Kissane, D. W., Grabsch, B., Clarke, D. M., Smith, G. C., Love, A. W., Bloch, S., . . . Li, Y. (2007). Supportive-expressive group therapy for women with metastatic breast cancer: Survival and psychosocial outcome from a randomized controlled trial. *Psycho-Oncology*, 16(4), 277-286. doi:10.1002/pon.1185

Lane, L. G., & Viney, L. L. (2005). The effects of personal construct group therapy on breast cancer survivors. *Journal of Consulting and Clinical Psychology*, 73(2), 284-292. doi:10.1037/0022-006x.73.2.284

Larson, M. R., Duberstein, P. R., Talbot, N. L., Caldwell, C., & Moynihan, J. A. (2000). A presurgical psychosocial intervention for breast cancer patients: psychological distress and the immune response. *Journal of Psychosomatic Research*, 48(2), 187-194. doi:10.1016/s0022-3999(99)00110-5

Lechner, S. C., Whitehead, N. E., Vargas, S., Annane, D. W., Robertson, B. R., Carver, C. S., . . . Antoni, M. H. (2014). Does a community-based stress management intervention affect psychological adaptation among underserved black breast cancer survivors? *Journal of the National Cancer Institute. Monographs National Cancer Institute*, 2014(50), 315-322. doi:10.1093/jncimonographs/lgu032

Lee, R., Lee, K. S., Oh, E.-G., & Kim, S. H. (2013). A Randomized Trial of Dyadic Peer Support Intervention for Newly Diagnosed Breast Cancer Patients in Korea. *Cancer Nursing*, 36(3), E15-E22. doi:10.1097/NCC.0b013e3182642d7c

Lengacher, C. A., Johnson-Mallard, V., Post-White, J., Moscoso, M. S., Jacobsen, P. B., Klein, T. W., . . . Kip, K. E. (2009). Randomized controlled trial of mindfulness-based stress reduction (MBSR) for survivors of breast cancer. *Psycho-Oncology*, 18(12), 1261-1272. doi:10.1002/pon.1529

Lewis, F. M., Brandt, P. A., Cochrane, B. B., Griffith, K. A., Grant, M., Haase, J. E., . . . Shands, M. E. (2015). The Enhancing Connections Program: A six-state randomized clinical trial of a cancer parenting program. *Journal of Consulting and Clinical Psychology*, 83(1), 12-23. doi:10.1037/a0038219

Maguire, P., Tait, A., Brooke, M., Thomas, C., & Sellwood, R. (1980). Effects of counselling on the psychiatric morbidity associated with mastectomy. *British Medical Journal*, 281(6253), 1454-1456.

Manne, S. L., Ostroff, J. S., Winkel, G., Fox, K., Grana, G., Miller, E., . . . Frazier, T. (2005). Couple-focused group intervention for women with early stage breast cancer. *Journal of Consulting and Clinical Psychology*, 73(4), 634-646. doi:10.1037/0022-006x.73.4.634

Manne, S. L., Siegel, S. D., Heckman, C. J., & Kashy, D. A. (2016). A randomized clinical trial of a supportive versus a skill-based couple-focused group intervention for breast cancer patients. *Journal of Consulting and Clinical Psychology*, 84(8), 668-681. doi:10.1037/ccp0000110

Marchioro, G., Azzarello, G., Checchin, F., Perale, M., Segati, R., Sampognaro, E., . . . Vinante, O. (1996). The impact of a psychological intervention on quality of life in non-metastatic breast cancer. *European Journal of Cancer,* 32A(9), 1612-1615. doi:10.1016/0959-8049(96)00134-7

Marcus, A. C., Garrett, K. M., Cella, D., Wenzel, L., Brady, M. J., Fairclough, D., . . . Flynn, P. J. (2010). Can telephone counseling post-treatment improve psychosocial outcomes among early stage breast cancer survivors? *Psycho-Oncology,* 19(9), 923-932. doi:10.1002/pon.1653

McArdle, J. M., George, W. D., McArdle, C. S., Smith, D. C., Moodie, A. R., Hughson, A. M., & Murray, G. D. (1996). Psychological support for patients undergoing breast cancer surgery: a randomised study. *BMJ*, 312(7034), 813-816.

Mens, M. G., Helgeson, V. S., Lembersky, B. C., Baum, A., & Scheier, M. F. (2016). Randomized psychosocial interventions for breast cancer: impact on life purpose. *Psycho-Oncology*, 25(6), 618-625. doi:10.1002/pon.3891

Miyashita, M. (2005). A randomized intervention study for breast cancer survivors in Japan. *Cancer Nursing*, 28(1), 70-78. doi:10.1097/00002820-200501000-00011

Nápoles, A. M., Ortíz, C., Santoyo-Olsson, J., Stewart, A. L., Gregorich, S., Lee, H. E., . . . Luce, J. (2015). Nuevo Amanecer: Results of a randomized controlled trial of a community-based, peer-delivered stress management intervention to improve quality of life in Latinas with breast cancer. *American Journal of Public Health*, 105(Suppl 3), e55-e63. doi:10.2105/AJPH.2015.302598

Naumann, F., Martin, E., Philpott, M., Smith, C., Groff, D., & Battaglini, C. (2012). Can counseling add value to an exercise intervention for improving quality of life in breast cancer survivors? A feasibility study. *The Journal of Supportive Oncology*, 10(5), 188-194. doi:10.1016/j.suponc.2011.09.004

Pelekasis, P., Zisi, G., Koumarianou, A., Marioli, A., Chrousos, G., Syrigos, K., & Darviri, C. (2016). Forming a stress management and health promotion program for women undergoing chemotherapy for breast cancer: A pilot randomized controlled trial. *Integrative Cancer Therapies*, 15(2), 165-174.

Qiu, J., Chen, W., Gao, X., Xu, Y., Tong, H., Yang, M., . . . Yang, M. (2013). A randomized controlled trial of group cognitive behavioral therapy for Chinese

breast cancer patients with major depression. *Journal of Psychosomatic Obstetrics and Gynecology*, 34(2), 60-67. doi:10.3109/0167482x.2013.766791

Rissanen, R., Nordin, K., Ahlgren, J., & Arving, C. (2015). A stepped care stress management intervention on cancer‐related traumatic stress symptoms among breast cancer patients−a randomized study in group vs. individual setting. *Psycho-Oncology*, 24(9), 1028-1035.

Samarel, N., Fawcett, J., & Tulman, L. (1997). Effect of support groups with coaching on adaptation to early stage breast cancer. *Research in Nursing & Health*, 20(1), 15-26.

Sandgren, A. K., & McCaul, K. D. (2003). Short-term effects of telephone therapy for breast cancer patients. *Health Psychology*, 22(3), 310-315. doi:10.1037/0278-6133.22.3.310

Sandgren, A. K., McCaul, K. D., King, B., O'Donnell, S., & Foreman, G. (2000). Telephone therapy for patients with breast cancer. *Oncology Nursing Forum*, 27(4), 683-688.

Savard, J., Ivers, H., Savard, M.-H., & Morin, C. M. (2014). Is a video-based cognitive behavioral therapy for insomnia as efficacious as a professionally administered treatment in breast cancer? Results of a randomized controlled trial. Sleep: *Journal of Sleep and Sleep Disorders Research*, 37(8), 1305-1314.

Savard, J., Simard, S., Giguere, I., Ivers, H., Morin, C. M., Maunsell, E., . . . Marceau, D. (2006). Randomized clinical trial on cognitive therapy for depression in women with metastatic breast cancer: psychological and immunological effects. *Palliative Support Care*, 4(3), 219-237.

Savard, J., Simard, S., Ivers, H., & Morin, C. M. (2005). Randomized study on the efficacy of cognitive-behavioral therapy for insomnia secondary to breast cancer, part I: Sleep and psychological effects. *Journal of Clinical Oncology*, *23*(25), 6083-6096.

Scheier, M. F., Helgeson, V. S., Schulz, R., Colvin, S., Berga, S., Bridges, M. W., . . . Pappert, W. S. (2005). Interventions to enhance physical and psychological functioning among younger women who are ending nonhormonal adjuvant treatment for early-stage breast cancer. *Journal of Clinical Oncology*, 23(19), 4298-4311.

Schnur, J. B., David, D., Kangas, M., Green, S., Bovbjerg, D. H., & Montgomery, G. H. (2009). A randomized trial of a cognitive-behavioral therapy and hypnosis intervention on positive and negative affect during breast cancer radiotherapy. *Journal of Clinical Psychology*, 65(4), 443-455. doi:10.1002/jclp.20559

Schover, L. R., Rhodes, M. M., Baum, G., Adams, J. H., Jenkins, R., Lewis, P., & Jackson, K. E. (2011). Sisters Peer Counseling in Reproductive Issues After Treatment (SPIRIT): A peer counseling program to improve reproductive health among African American breast cancer survivors. *Cancer*, 117(21), 4983-4992. doi:10.1002/cncr.26139

Spiegel, D., Bloom, J. R., & Yalom, I. (1981). Group support for patients with metastatic cancer: A randomized prospective outcome study. *Archives of General Psychiatry,* 38(5), 527-533.

Stanton, A. L., Ganz, P. A., Kwan, L., Meyerowitz, B. E., Bower, J. E., Krupnick, J. L., . . . Belin, T. R. (2005). Outcomes from the moving beyond cancer psychoeducational, randomized, controlled trial with breast cancer patients. *Journal of Clinical Oncology*, 23(25), 6009-6018. doi:10.1200/JCO.2005.09.101

Steven, J., Simpson, A., Carlson, L. E., & Trew, M. E. (2001). Effect of group therapy for breast cancer on healthcare utilization. *Cancer Practice,* 9(1), 19-26. doi:10.1046/j.1523-5394.2001.91005.x

Taylor, E. E., & Ingleton, C. (2003). Hypnotherapy and cognitive-behaviour therapy in cancer care: the patients' view. *European Journal of Cancer Care*, 12(2), 137-142. doi:10.1046/j.1365-2354.2003.00371.x

Vos, P. J., Visser, A. P., Garssen, B., Duivenvoorden, H. J., & de Haes, H. C. J. M. (2007). Effectiveness of group psychotherapy compared to social support groups in patients with primary, non-metastatic breast cancer. *Journal of Psychosocial Oncology*, 25(4), 37-60. doi:10.1300/J077v25n04_03

Wurtzen, H., Dalton, S. O., Elsass, P., Sumbundu, A. D., Steding-Jensen, M., Karlsen, R. V., . . . Johansen, C. (2013). Mindfulness significantly reduces self-reported levels of anxiety and depression: results of a randomised controlled trial among 336 Danish women treated for stage I-III breast cancer. *European Journal of Cancer*, 49(6), 1365-1373. doi:10.1016/j.ejca.2012.10.030

Yates, P., Aranda, S., Hargraves, M., Mirolo, B., Clavarino, A., McLachlan, S., & Skerman, H. (2005). Randomized controlled trial of an educational intervention for managing fatigue in women receiving adjuvant chemotherapy for early-stage breast cancer. *Journal of Clinical Oncology*, 23(25), 6027-6036.

Yvonne Wengström, C. H., Hans Strander, Christina Forsberg, Yvonne. (1999). Effects of a nursing intervention on subjective distress, side effects and quality of life of breast cancer patients receiving curative radiation therapy: A randomized study. *Acta Oncologica*, 38(6), 763-770.

Zgâia, A., Pop, F., Achimaș-Cadariu, P., Vlad, C., Rogobete, A., Lisencu, C., . . . Irimie, A. (2016). The impact of relaxation technique and pre-operative psychological counselling on pain, analgesic consumption and psychological symptoms on patients scheduled for breast cancer surgery – A randomized clinical study. *Journal of Evidence-Based Psychotherapies*, 16(2), 205-220.

Appendix 6:

References of the RCTs included (study 3)

References of Included RCTs

Andersen, B. L., Farrar, W. B., Golden-Kreutz, D., Emery, C. F., Glaser, R., Crespin, T., & Carson III, W. E. (2007). Distress reduction from a psychological intervention contributes to improved health for cancer patients. *Brain, Behavior, and Immunity*, 21(7), 953-961.

Andersen, B. L., Farrar, W. B., Golden-Kreutz, D. M., Glaser, R., Emery, C. F., Crespin, T. R., . . . Carson, W. E. (2004). Psychological, behavioral, and immune changes after a psychological intervention: A clinical trial. Journal of *Clinical Oncology*, 22(17), 3570-3580. doi: 10.1200/jco.2004.06.030

Antoni, M. H., Lehman, J. M., Kilbourn, K. M., Boyers, A. E., Culver, J. L., Alferi, S. M., . . . Harris, S. D. (2001). Cognitive-behavioral stress management intervention decreases the prevalence of depression and enhances benefit finding among women under treatment for early-stage breast cancer. *Health Psychology*, 20(1), 20-32.

Beutel, M. E., Weissflog, G., Leuteritz, K., Wiltink, J., Haselbacher, A., Ruckes, C., . . . Braehler, E. (2014). Efficacy of short-term psychodynamic psychotherapy (STPP) with depressed breast cancer patients: results of a randomized controlled multicenter trial. *Annals of Oncology*, 25(2), 378-384. doi: 10.1093/annonc/mdt526

Bredal, I. S., Karesen, R., Smeby, N. A., Espe, R., Sorensen, E. M., Amundsen, M., . . . Ekeberg, O. (2014). Effects of a Psychoeducational Versus a Support Group Intervention in Patients with Early-Stage Breast Cancer Results of a Randomized Controlled Trial. *Cancer Nursing*, 37(3), 198-207. doi: 10.1097/NCC.0b013e31829879a3

Carlson, L. E., Doll, R., Stephen, J., Faris, P., Tamagawa, R., Drysdale, E., & Speca, M. (2013). Randomized controlled trial of mindfulness-based cancer recovery versus supportive expressive group therapy for distressed survivors of breast cancer. *Journal of Clinical Oncology*, 31(25), 3119-3126.

Classen, C. C., Kraemer, H. C., Blasey, C., Giese-Davis, J., Koopman, C., Palesh, O. G., . . . Spiegel, D. (2008). Supportive-expressive group therapy for primary breast cancer patients: a randomized prospective multicenter trial. *Psycho-Oncology*, 17(5), 438-447. doi: 10.1002/pon.1280

Desautels, C., Savard, J., Ivers, H., Savard, M.-H., & Caplette-Gingras, A. (2018). Treatment of depressive symptoms in patients with breast cancer: A randomized controlled trial comparing cognitive therapy and bright light therapy. *Health Psychology*, 37(1), 1-13.

Dirksen, S. R., & Epstein, D. R. (2008). Efficacy of an insomnia intervention on fatigue, mood and quality of life in breast cancer survivors. *Journal of Advanced Nursing*, 61(6), 664-675. doi: 10.1111/j.1365-2648.2007.04560.x

Graves, K. D., Carter, C. L., Anderson, E. S., & Winett, R. A. (2003). Quality of life pilot intervention for breast cancer patients: use of social cognitive theory. *Palliative Support Care*, 1(2), 121-134.

Groarke, A., Curtis, R., & Kerin, M. (2013). Cognitive-behavioural stress management enhances adjustment in women with breast cancer. *British Journal of Health Psychology*, 18(3), 623-641. doi: 10.1111/bjhp.12009

Ho, R. T., Fong, T. C., Lo, P. H., Ho, S. M., Lee, P. W., Leung, P. P., . . . Chan, C. L. (2016). Randomized controlled trial of supportive-expressive group therapy and body-mind-spirit intervention for Chinese non-metastatic breast cancer patients. *Support Care Cancer*, 24(12), 4929-4937. doi: 10.1007/s00520-016-3350-8

Lechner, S. C., Whitehead, N. E., Vargas, S., Annane, D. W., Robertson, B. R., Carver, C. S., . . . Antoni, M. H. (2014). Does a community-based stress management intervention affect psychological adaptation among underserved black breast cancer survivors? *Journal of the National Cancer Institute. Monographs National Cancer Institute*, 2014(50), 315-322. doi: 10.1093/jncimonographs/lgu032

Merckaert, I., Lewis, F., Delevallez, F., Herman, S., Caillier, M., Delvaux, N., . . . Razavi, D. (2017). Improving anxiety regulation in patients with breast cancer at the beginning of the survivorship period: a randomized clinical trial comparing the benefits of single-component and multiple-component group interventions. Psycho-Oncology, 26(8), 1147-1154. doi: 10.1002/pon.4294

Sandgren, A. K., & McCaul, K. D. (2003). Short-term effects of telephone therapy for breast cancer patients. Health Psychology, 22(3), 310-315. doi: 10.1037/0278-6133.22.3.310

Sandgren, A. K., & McCaul, K. D. (2007). Long-term telephone therapy outcomes for breast cancer patients. Psycho-Oncology: *Journal of the Psychological, Social and Behavioral Dimensions of Cance*r, 16(1), 38-47.

Savard, J., Simard, S., Giguere, I., Ivers, H., Morin, C. M., Maunsell, E., . . . Marceau, D. (2006). Randomized clinical trial on cognitive therapy for depression in women with metastatic breast cancer: psychological and immunological effects. *Palliative Support Care*, 4(3), 219-237.

Savard, J., Simard, S., Ivers, H., & Morin, C. M. (2005). Randomized study on the efficacy of cognitive-behavioral therapy for insomnia secondary to breast cancer, part I: Sleep and psychological effects. *Journal of Clinical Oncology*, 23(25), 6083-6096.

Stanton, A. L., Ganz, P. A., Kwan, L., Meyerowitz, B. E., Bower, J. E., Krupnick, J. L., . . . Belin, T. R. (2005). Outcomes from the moving beyond cancer psychoeducational, randomized, controlled trial with breast cancer patients. *Journal of Clinical Oncology*, 23(25), 6009-6018. doi: 10.1200/JCO.2005.09.101

Appendix 7:

Funnel plots of effect sizes against standard error data from eligible published trials (study 3)

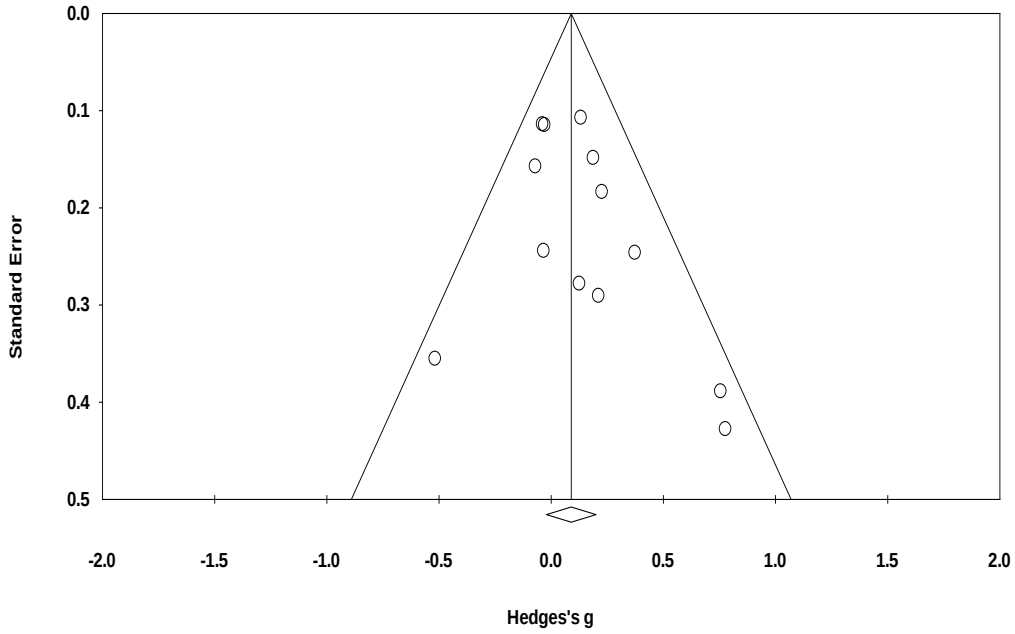


Figure 7A: Funnel plot of effect sizes against standard error for anxiety using data from eligible published trials at post-treatment

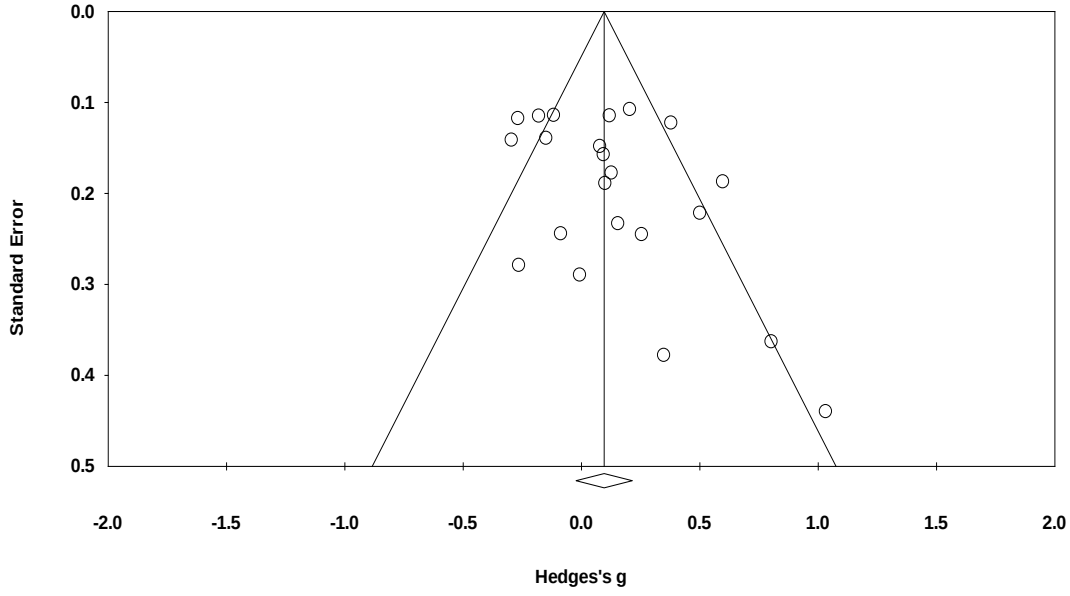


Figure 7B: Funnel plot of effect sizes against standard error for depression using data from eligible published trials at post-treatment

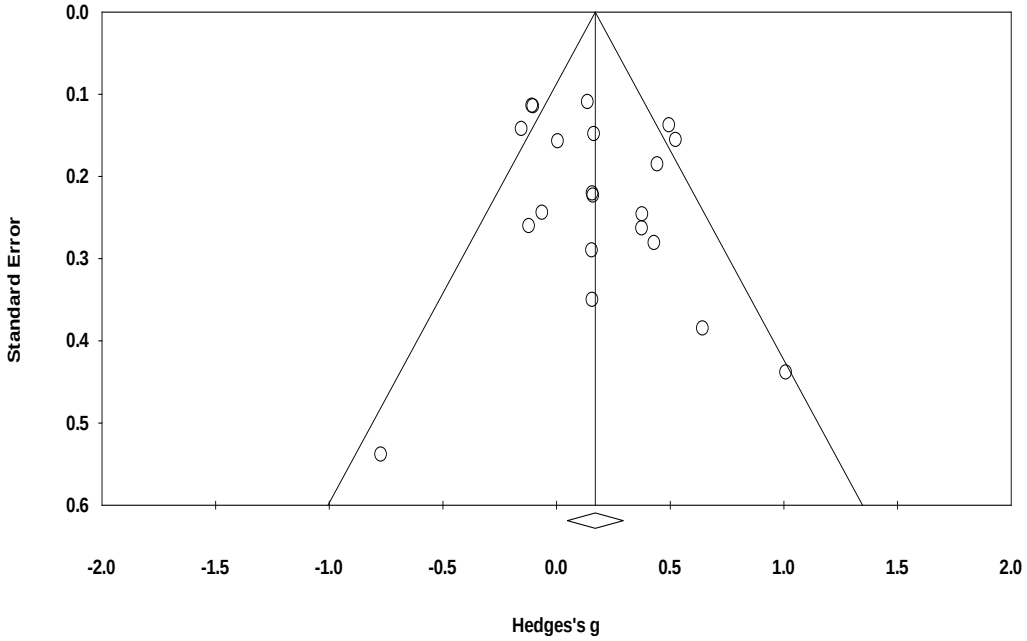
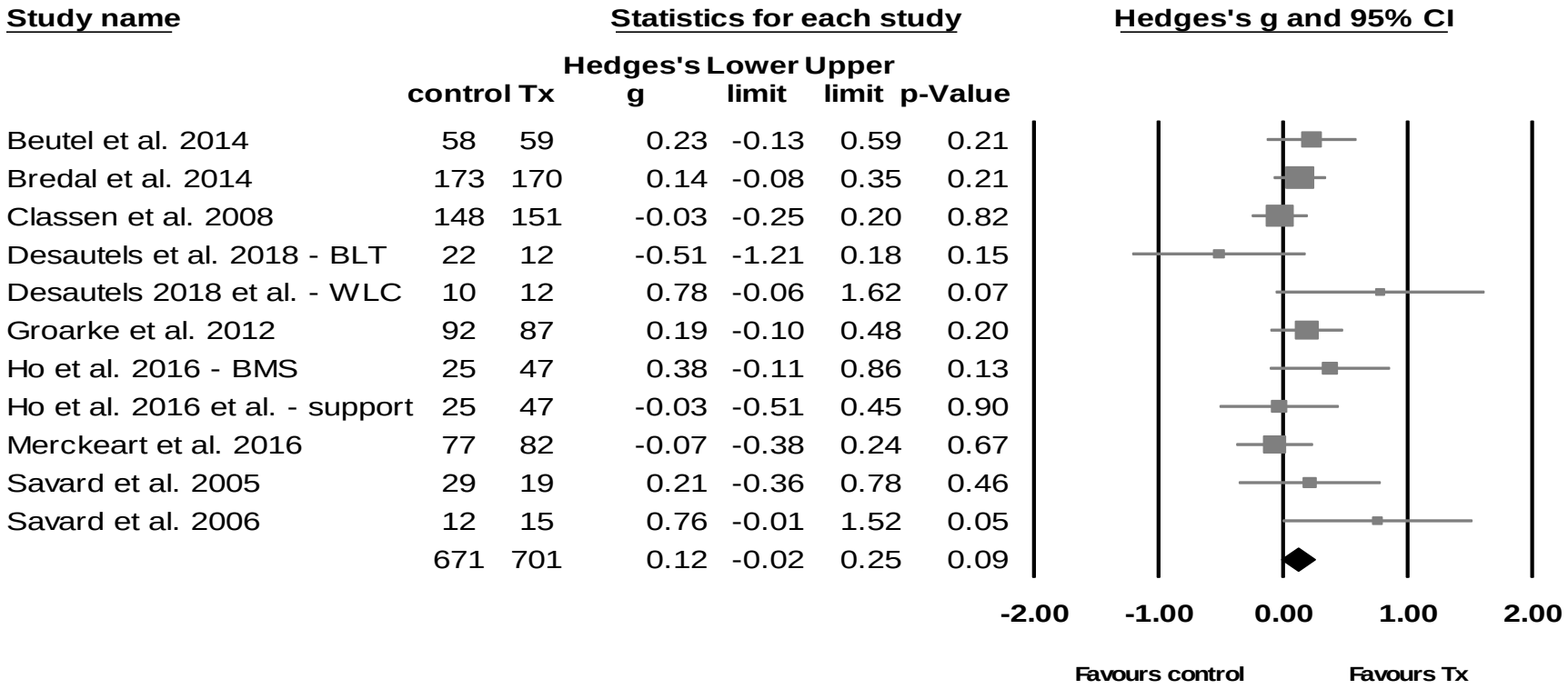


Figure 7C: Funnel plot of effect sizes against standard error for general distress using data from eligible published trials at post-treatment

Appendix 8:

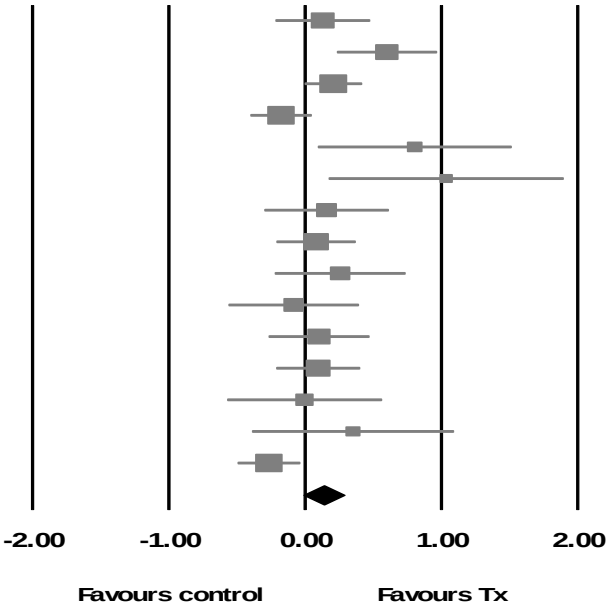
Treatment effects at post-treatment and follow-up for anxiety, depression and general distress in the total sample based on effect sizes (study 3)



Note. BLT = bright light therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy

Figure 8A: Forest plot of effect sizes for anxiety at post treatment in the total sample

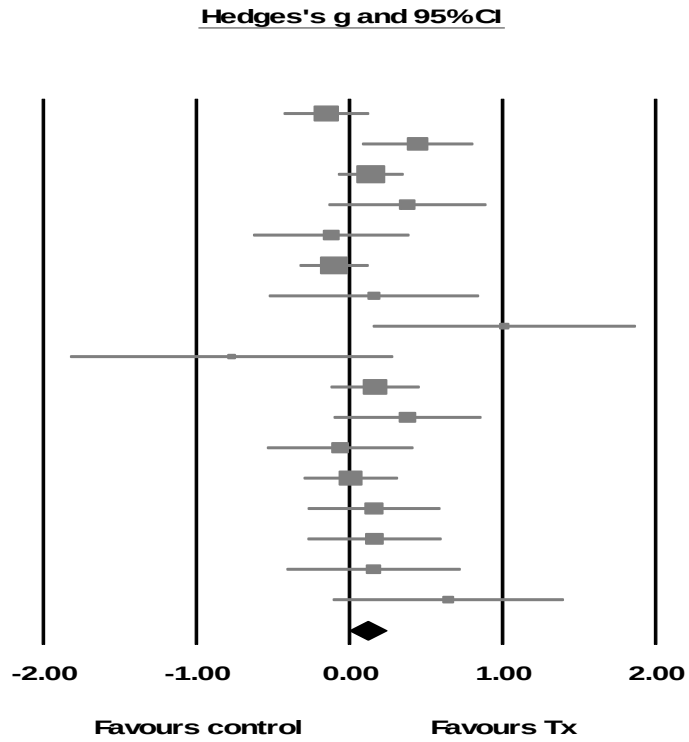
Study name	control	Tx	Hedges's g	Lower limit	Upper limit	p-Value
Antoni et al. 2001	72	55	0.13	-0.22	0.48	0.47
Beutel et al. 2014	58	59	0.60	0.23	0.97	0.00
Bredal et al. 2014	163	179	0.21	-0.01	0.42	0.06
Classen et al. 2008	148	151	-0.18	-0.40	0.05	0.13
Desautels et al. 2018 - BLT	22	12	0.80	0.09	1.52	0.03
Desautels et al. 2018 - WLC	10	12	1.03	0.17	1.90	0.02
Dirksen et al. 2008	38	34	0.16	-0.30	0.62	0.50
Groarke et al. 2013	92	87	0.08	-0.21	0.37	0.59
Ho et al. 2016 - BMS	25	47	0.26	-0.22	0.74	0.30
Ho et al. 2016 - support	25	47	-0.08	-0.56	0.40	0.73
Lechner et al. 2014	57	53	0.10	-0.27	0.47	0.59
Merckeart et al. 2016	77	82	0.10	-0.21	0.41	0.54
Savard et al. 2005	29	19	-0.00	-0.57	0.56	0.99
Savard et al. 2006	12	15	0.35	-0.39	1.09	0.35
Stanton et al. 2005	142	145	-0.27	-0.50	-0.03	0.02
	970	997	0.13	-0.01	0.28	0.08



Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy

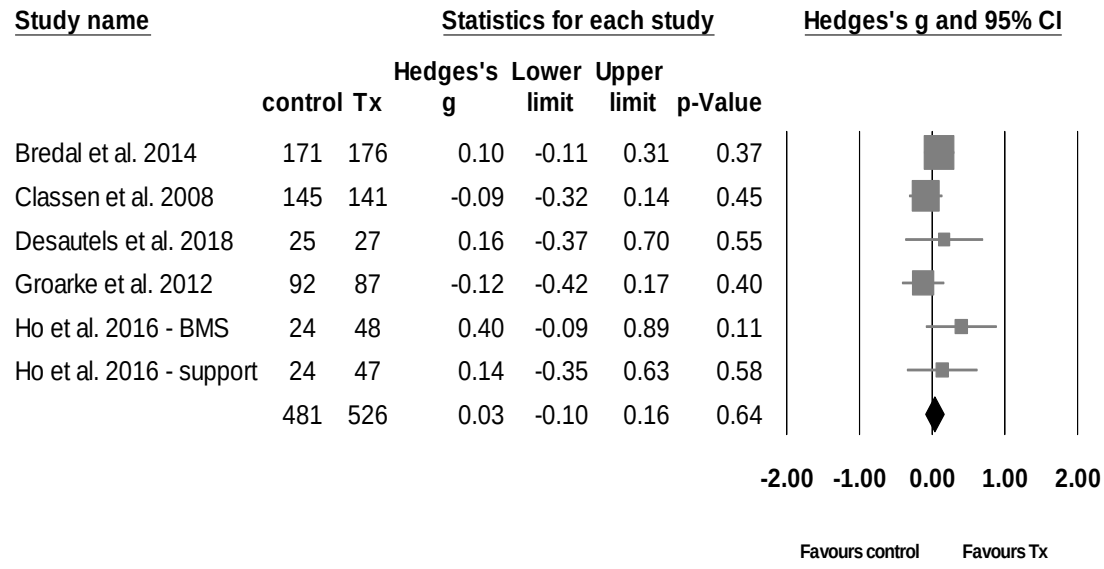
Figure 8B: Forest plot of effect sizes for depression at post treatment in the total sample

Study name	Sample size		Statistics for each study			
	control	Tx	Hedges's g	Lower limit	Upper limit	p-Value
Andersen et al. 2004	90	106	-0.15	-0.43	0.13	0.29
Beutel et al. 2014	58	59	0.45	0.08	0.81	0.02
Bredal et al. 2014	160	169	0.14	-0.08	0.36	0.21
Carlson et al. 2013 - MBT	18	69	0.38	-0.14	0.90	0.15
Carlson et al. 2013 - support	18	73	-0.12	-0.63	0.39	0.65
Classen et al. 2008	148	151	-0.10	-0.33	0.13	0.38
Desautels et al. 2018 - BLT	22	12	0.16	-0.53	0.85	0.65
Desautels et al. 2018 - WLC	10	12	1.01	0.15	1.87	0.02
Graves et al. 2003	6	7	-0.77	-1.83	0.29	0.15
Groarke et al. 2013	92	87	0.17	-0.13	0.46	0.26
Ho et al. 2016 - BMS	25	47	0.38	-0.10	0.86	0.12
Ho et al. 2016 - support	25	47	-0.06	-0.54	0.42	0.80
Merckeart et al. 2016	77	82	0.01	-0.30	0.32	0.96
Sandgren et al. 2003 - EE	26	90	0.16	-0.27	0.59	0.47
Sandgren et al. 2003 - PsyEd	26	81	0.16	-0.28	0.60	0.47
Savard et al. 2005	29	19	0.16	-0.41	0.73	0.59
Savard et al. 2006	12	15	0.65	-0.11	1.40	0.09
	842	1126	0.11	-0.00	0.23	0.06



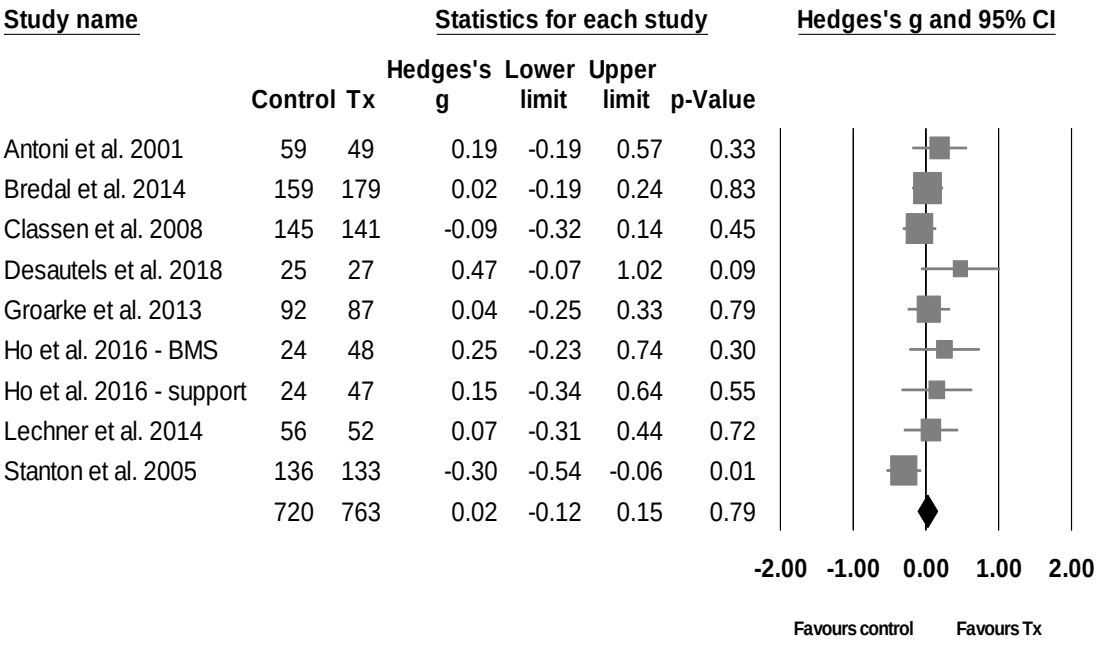
Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy; EE = emotional expression; PsyEd = psychoeducation

Figure 8C: Forest plot of effect sizes for general distress at post treatment in the total sample



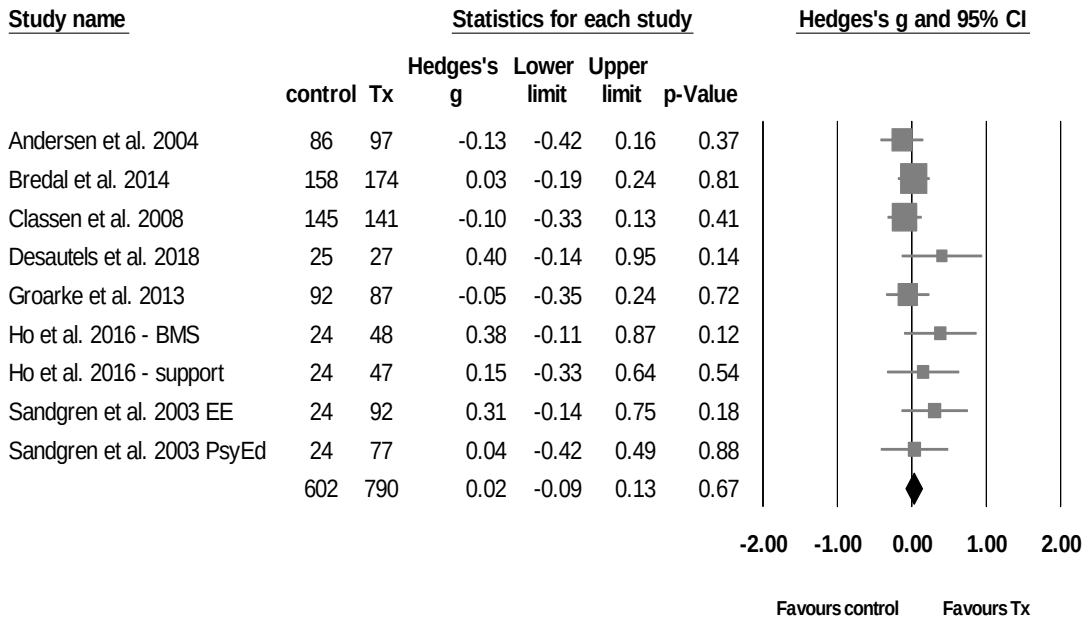
Note; BMS = body mind spirit; support = supportive therapy

Figure 8D: Forest plot of effect sizes for anxiety at follow-up in the total sample



Note. BMS = body mind spirit; support = supportive therapy

Figure 8E: Forest plot of effect sizes for depression at follow-up in the total sample

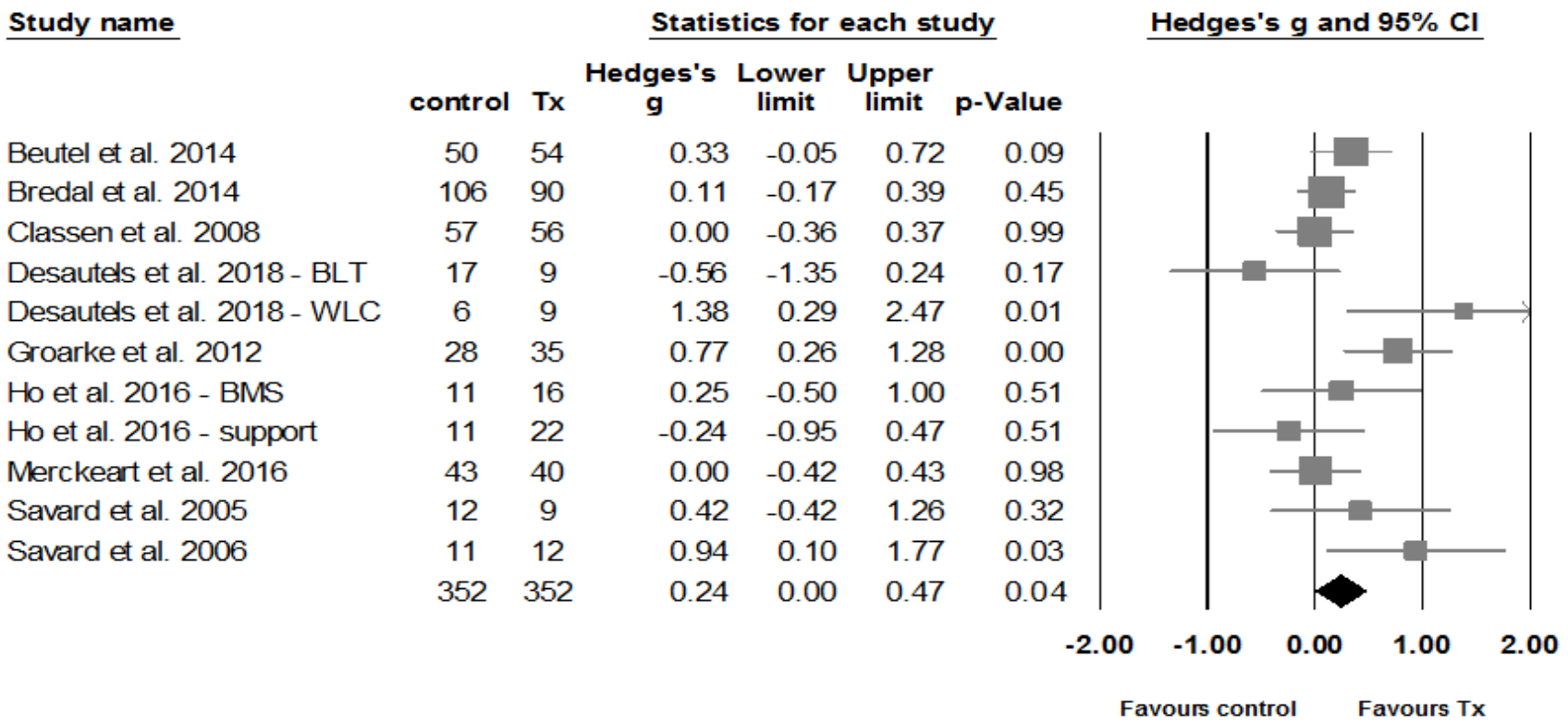


Note. BMS = body mind spirit; support = supportive therapy; EE = emotional expression; PsyEd = psychoeducation

Figure 8F: Forest plot of effect sizes for general distress at follow-up in the total sample

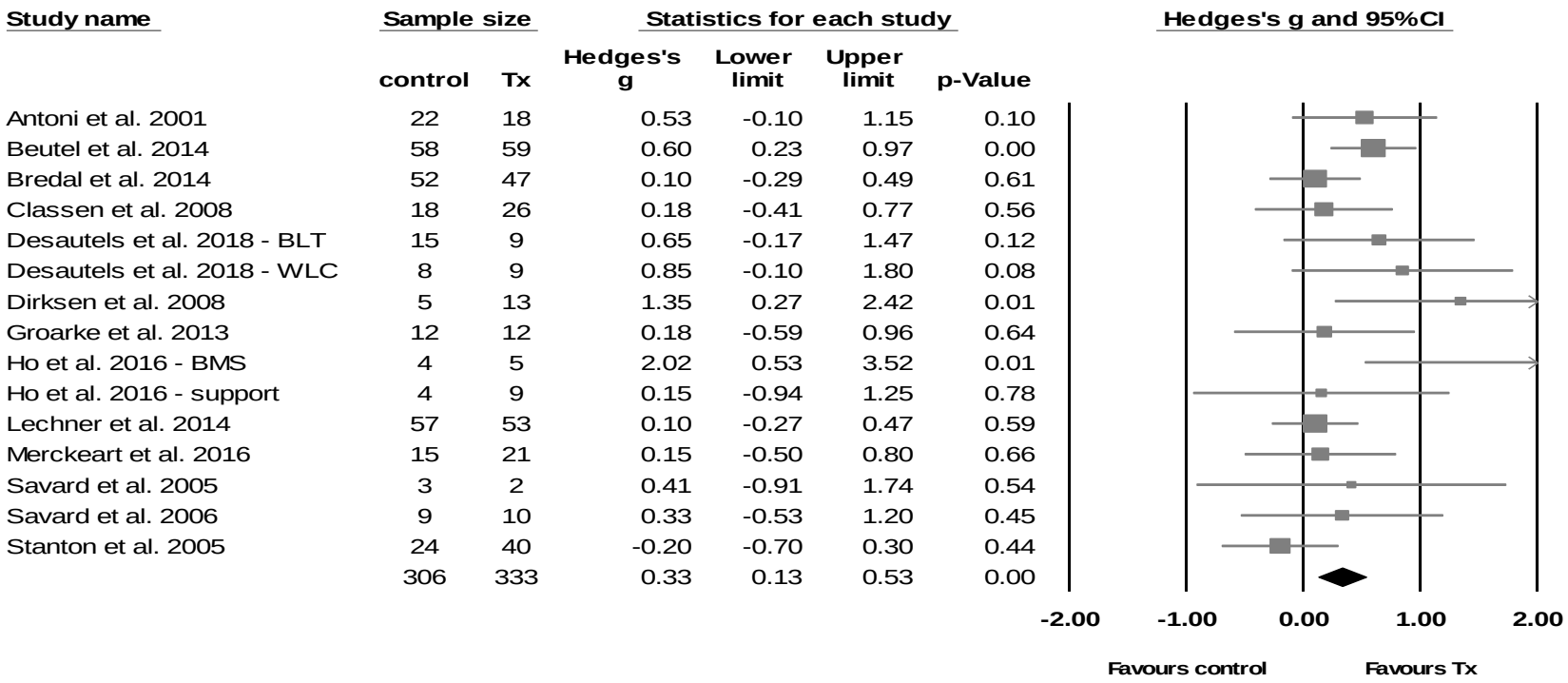
Appendix 9:

Treatment effects at post-treatment for anxiety, depression and general distress in the distressed sub-sample based on effect sizes (study 3)



Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy

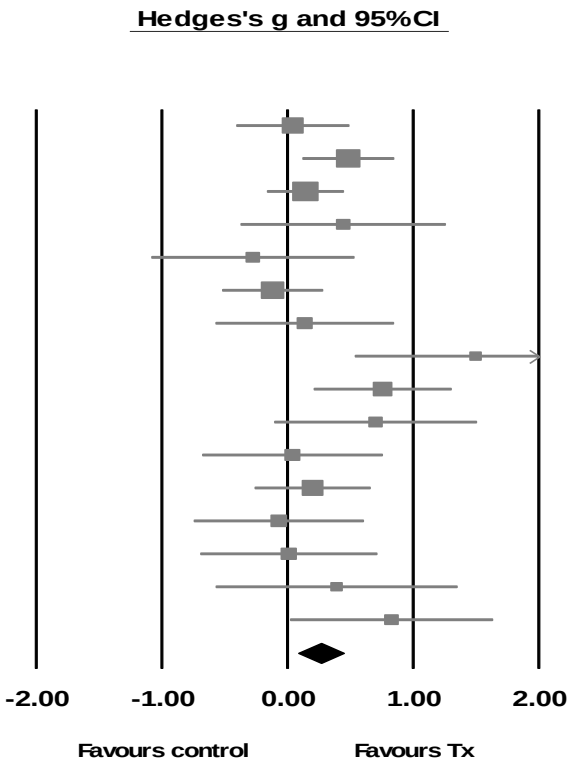
Figure 9A: Forest plot of effect sizes for anxiety at post treatment in the distressed sub-sample



Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy

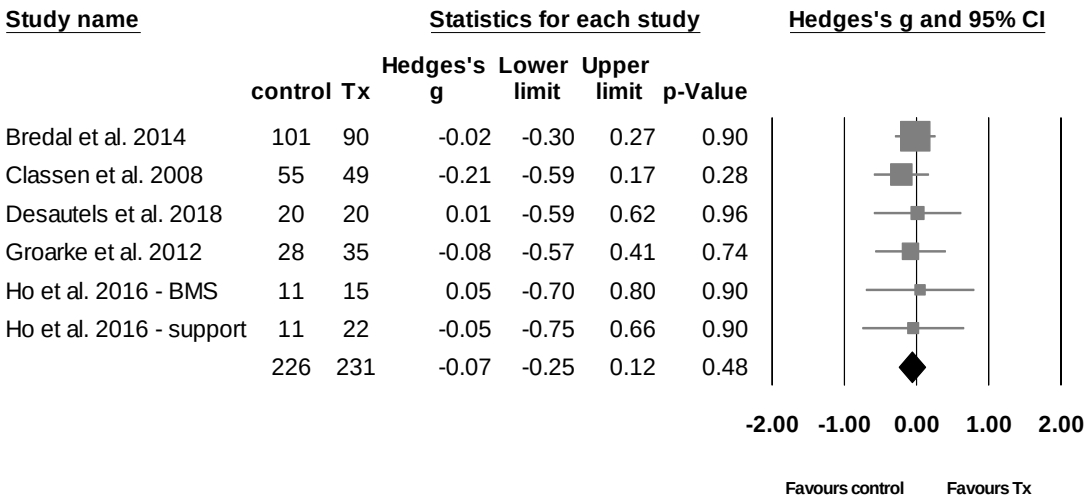
Figure 9B: Forest plot of effect sizes for depression at post treatment in the distressed sub-sample

Study name	Sample size		Statistics for each study			
	control	Tx	Hedges's g	Lower limit	Upper limit	p-Value
Andersen et al. 2004	30	48	0.04	-0.41	0.49	0.86
Beutel et al. 2014	57	59	0.48	0.12	0.85	0.01
Bredal et al. 2014	80	81	0.14	-0.17	0.45	0.36
Carlson et al. 2013 - MBT	7	27	0.44	-0.38	1.26	0.29
Carlson et al. 2013 - support	7	29	-0.28	-1.08	0.53	0.50
Classen et al. 2008	44	49	-0.12	-0.52	0.29	0.57
Desautels et al. 2018 - BLT	21	11	0.14	-0.57	0.85	0.71
Desautels et al. 2018 - WLC	9	11	1.50	0.53	2.46	0.00
Groarke et al. 2013	25	28	0.76	0.21	1.31	0.01
Ho et al. 2016 - BMS	10	14	0.70	-0.11	1.51	0.09
Ho et al. 2016 - support	10	24	0.04	-0.68	0.76	0.92
Merckeart et al. 2016	32	39	0.20	-0.26	0.66	0.40
Sandgren et al. 2003 - EE	11	30	-0.07	-0.75	0.61	0.84
Sandgren et al. 2003 - PsyEd	11	22	0.01	-0.70	0.72	0.98
Savard et al. 2005	7	8	0.39	-0.57	1.36	0.43
Savard et al. 2006	11	13	0.83	0.02	1.64	0.05
	372	493	0.26	0.08	0.44	0.00



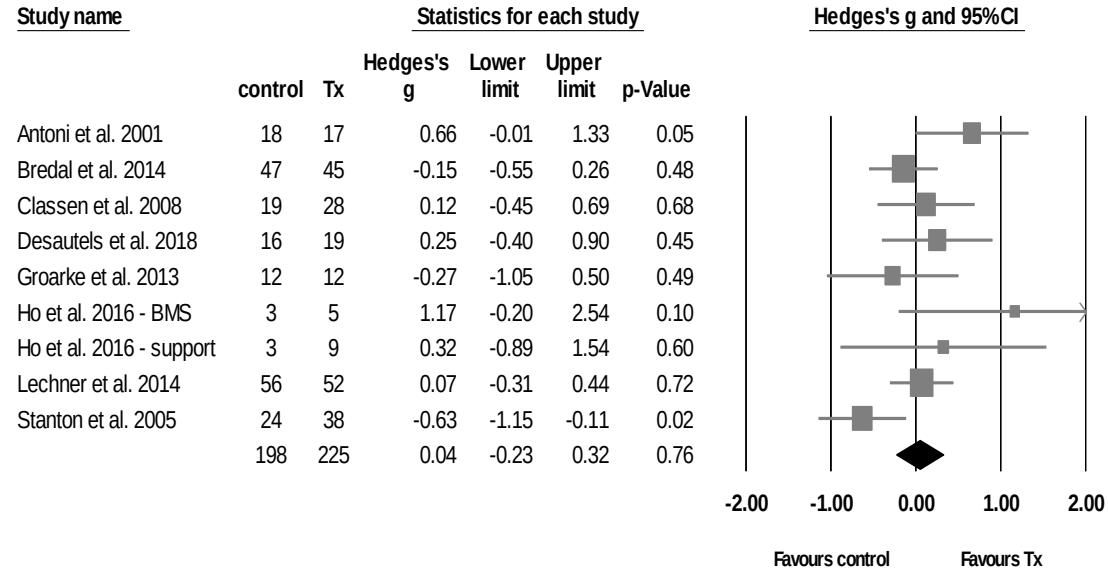
Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy; EE = emotional expression**;** PsyEd = psychoeducation

Figure 9C: Forest plot of effect sizes for general distress at post treatment in the distressed sub-sample



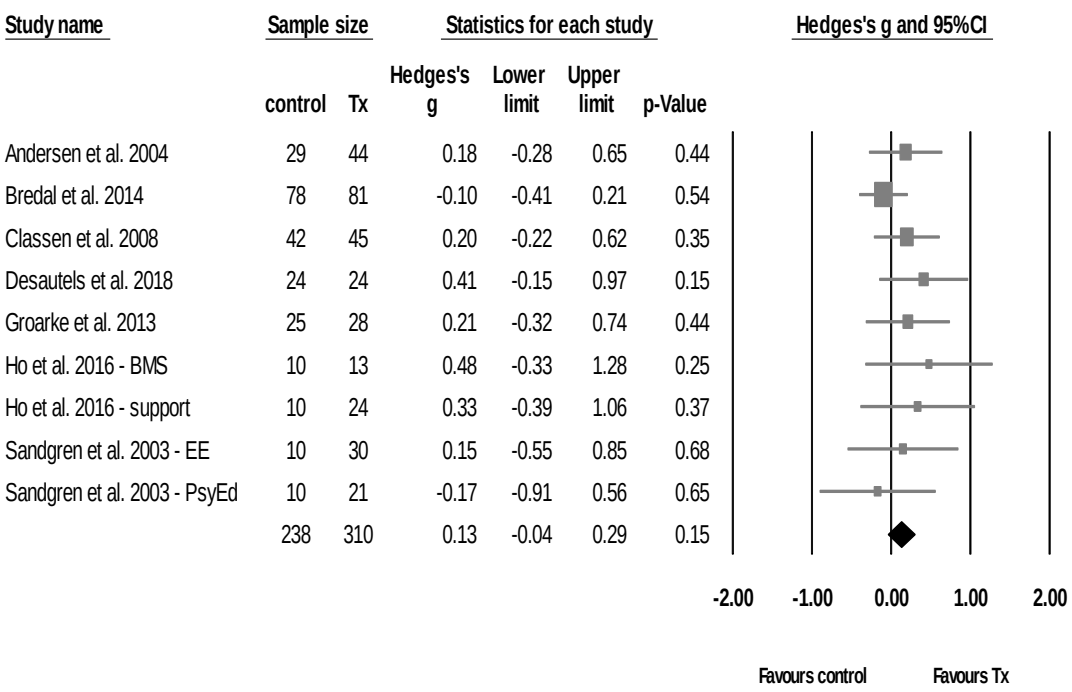
Note. BMS = body mind spirit; support = supportive therapy

Figure 9D: Forest plot of effect sizes for anxiety at follow-up in the distressed sub-sample



Note. BMS = body mind spirit; support = supportive therapy

Figure 9E: Forest plot of effect sizes for depression at follow-up in the distressed sub-sample



Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy; EE = emotional expression; PsyEd = psychoeducation

Figure 9F: Forest plot of effect sizes for general distress at follow-up in the distressed sub-sample

Appendix 10:

Classification of patients at post-treatment and follow-up according to Jacobson's second criterion (RCI) for anxiety, depression and general distress in the total sample (study 3)

Table 10A: Classification of patients at post-treatment and follow-up according to Jacobson's second criterion (RCI) for anxiety in the total sample

			Post-treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
Beutel et al. 2014	Treatment	Other - Psychodynamic therapy	59	10	41	49	Follow-up data not reported			
	95% CI			(4-21)	(28-54)	(36-63)				
	Control	Treatment as usual - Offered referral to GP for psychological or pharmacological treatment	58	9	60	31				
	95% CI			(3-19)	(47-73)	(20-45)				
Bredal et al. 2014	Treatment	Psychoeducation	170	18	45	37	176	9	36	55
	95% CI			(12-24)	(38-53)	(30-45)		(5-14)	(29-44)	(47-62)
	Control	Treatment as usual - Nurse-led support group	173	17	49	34	171	10	41	49
	95% CI			(12-24)	(41-56)	(27-42)		(6-15)	(33-49)	(41-57)
Classen et al. 2008	Treatment	Supportive therapy + education material	151	16	66	19	141	16	60	24
	95% CI			(10-23)	(57-73)	(13-26)		(11-23)	(51-68)	(17-32)
	Control	Education material	148	15	67	18	145	12	65	23
	95% CI			(10-22)	(59-74)	(12-25)		(7-18)	(56-73)	(27-31)

			Post-treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
Desautels et al. 2018	Treatment	Cognitive behavioural therapy	25	0	76	24	27	4	30	67
	95% CI			(0-14)	(55-91)	(9-45)		(0-19)	(14-50)	(46-83)
	Control	Bright light therapy	22	9	14	77	25	4	32	64
	95% CI			(1-29)	(3-35)	(55-92)		(0-20)	(15-54)	(43-82)
	Control	Wait-list control	10	10	60	30	Not applicable			
	95% CI			(0-45)	(26-88)	(7-65)				
Groarke et al. 2013	Treatment	Cognitive behavioural therapy	87	8	59	33	87	17	53	30
	95% CI			(3-16)	(48-69)	(24-44)		(10-27)	(42-64)	(21-41)
	Control	Treatment as usual - Support from oncology nurses	92	23	65	12	92	18	61	21
	95% CI			(15-33)	(55-75)	(6-20)		(11-28)	(50-71)	(13-30)
Ho et al. 2016	Treatment	Other - Body-mind-spirit	45	11	67	22	46	17	63	20
	95% CI			(4-24)	(51-80)	(11-37)		(8-31)	(48-77)	(9-34)
	Treatment	Supportive therapy	47	19	66	15	47	15	70	15
	95% CI			(9-33)	(51-79)	(6-28)		(6-28)	(55-83)	(6-28)
	Control	Self-help support group	51	14	67	20	48	19	54	27

			Post-treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
	95% CI			(6-26)	(52-79)	(10-33)		(9-33)	(39-69)	(15-42)
Merckaert et al. 2016	Treatment	Cognitive behavioural therapy	82	11	62	27	Follow-up data not reported			
	95% CI			(5-20)	(51-73)	(18-38)				
	Control	Treatment as usual - Peer support group	77	5	71	23				
	95% CI			(1-13)	(60-81)	(14-34)				
Savard et al. 2005	Treatment	Cognitive behavioural therapy	19	0	74	26	Not applicable			
	95% CI			(0-18)	(49-91)	(9-51)				
	Control	Wait-list control	29	24	66	10				
	95% CI			(10-44)	(46-82)	(2-27)				
Savard et al. 2006	Treatment	Cognitive behavioural therapy	15	0	27	73	Not applicable			
	95% CI			(0-22)	(8-55)	(45-92)				
	Control	Wait-list control	12	17	50	33				
	95% CI			(2-48)	(21-79)	(10-65)				

Note. n = number of patients; RCI = reliable change index at P<0.05; CI = confidence interval

Table 10B: Classification of patients at post-treatment and follow-up according to Jacobson's second criterion (RCI) for depression in the total sample

			Post treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
Antoni et al. 2001	Treatment	Cognitive behavioural therapy	54	13	70	17	48	6	65	29
	95% CI			(5-25)	(56-82)	(8-29)		(1-17)	(49-78)	(17-44)
	Control	Condensed version of active treatment	72	8	78	14	59	7	78	15
	95% CI			(3-17)	(66-87)	(7-24)		(2-16)	(65-88)	(7-27)
Beutel et al. 2014	Treatment	Other - Psychodynamic therapy	59	3	24	73	Follow-up data not reported			
	95% CI			(0-12)	(14-37)	(60-84)				
	Control	Treatment as usual - Offered referral to GP for psychological or pharmacological treatment	58	17	41	41				
	95% CI			(9-29)	(29-55)	(29-55)				
Bredal et al. 2014	Treatment	Psychoeducation	177	19	67	15	177	12	56	32
	95% CI			(13-25)	(59-74)	(10-21)		(7-18)	(48-63)	(25-40)

			Post treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
	Control	Treatment as usual - Nurse-led support group	156	17	65	18	152	12	45	43
	95% CI			(12-24)	(57-72)	(12-25)		(7-18)	(37-53)	(35-52)
Classen et al. 2008	Treatment	Supportive therapy+ education material	151	15	67	18	141	9	70	22
	95% CI			(10-22)	(59-74)	(12-25)		(4-14)	(61-77)	(15-30)
	Control	Education material	148	11	71	18	145	12	69	19
	95% CI			(7-18)	(63-78)	(12-25)		(8-19)	(61-76)	(13-26)
Desautels et al. 2018	Treatment	Cognitive behavioural therapy	25	0	48	52	27	4	22	74
	95% CI			(0-14)	(28-69)	(31-72)		(0-19)	(9-42)	(54-89)
	Control	Bright light therapy	22	9	36	55	25	8	32	60
	95% CI			(1-29)	(17-59)	(32-76)		(0-26)	(15-54)	(39-79)
	Control	Wait-list control	10	0	70	30	Not applicable			
	95% CI			(0-31)	(35-93)	(7-65)				
Dirksen et al. 2008	Treatment	Cognitive behavioural therapy	34	3	76	21	Follow-up data not reported			

			Post treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
	95% CI			(0-15)	(59-89)	(9-38)				
	Control	Educational components of active treatment	38	3	87	11				
	95% CI			(0-14)	(72-96)	(3-25)				
Groarke et al. 2013	Treatment	Cognitive behavioural therapy	87	13	68	20	87	11	70	18
	95% CI			(6-21)	(57-77)	(12-29)		(6-20)	(59-79)	(11-28)
	Control	Treatment as usual - Support from oncology nurses	92	12	78	10	92	13	73	14
	95% CI			(6-20)	(68-86)	(5-18)		(7-22)	(63-82)	(8-23)
Ho et al. 2016	Treatment	Other - Body-mind-spirit intervention	45	11	76	13	46	7	80	13
	95% CI			(4-24)	(60-87)	(5-27)		(1-18)	(66-91)	(5-26)
	Treatment	Supportive therapy	47	15	72	13	47	9	79	13
	95% CI			(6-28)	(57-84)	(5-26)		(2-20)	(64-89)	(5-26)
	Control	Self-help support group	51	16	71	14	48	19	67	15

			Post treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
	95% CI			(7-29)	(56-83)	(6-26)		(9-33)	(52-80)	(6-28)
Lechner et al. 2014	Treatment	Cognitive behavioural therapy	53	8	75	17	52	4	75	21
	95% CI			(2-18)	(62-86)	(8-30)		(0-13)	(61-86)	(11-35)
	Control	Educational information delivered by therapist	57	12	67	21	56	9	66	25
	95% CI			(5-24)	(53-79)	(11-34)		(3-20)	(52-78)	(14-38)
Merckaert et al. 2016	Treatment	Cognitive behavioural therapy	82	6	65	29	Follow-up data not reported			
	95% CI			(2-14)	(53-75)	(20-40)				
	Control	Treatment as usual - Peer support group	77	14	70	16				
	95% CI			(7-24)	(59-80)	(8-26)				
Savard et al. 2005	Treatment	Cognitive behavioural therapy	19	0	74	26	Not applicable			
	95% CI			(0-18)	(49-91)	(9-51)				
	Control	Wait-list control	29	7	83	10				

			Post treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
	95% CI			(0-23)	(64-94)	(2-27)				
Savard et al. 2006	Treatment	Cognitive behavioural therapy	15	0	27	73	Not applicable			
	95% CI			(0-22)	(8-55)	(45-92)				
	Control	Wait-list control	12	0	58	42				
	95% CI			(0-26)	(28-85)	(15-72)				
Stanton et al. 2005	Treatment	Psychoeducation + educational material	144	9	79	12	132	6	83	11
	95% CI			(5-15)	(72-85)	(7-18)		(3-12)	(76-89)	(6-17)
	Control	Education material	142	8	83	8	136	8	82	10
	95% CI			(4-14)	(76-89)	(4-14)		(4-14)	(75-88)	(5-16)

Note. n = number of patients; RCI = reliable change index at P<0.05; CI = confidence interval

Table 10C: Classification of patients at post-treatment and follow-up according to Jacobson's second criterion (RCI) for general distress in the total sample

			Post treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
Andersen et al. 2004	Treatment	Other - Biobehavioral intervention	106	12	51	37	97	10	54	36
	95% CI			(7-20)	(41-61)	(28-47)		(5-18)	(43-64)	(27-46)
	Control	Assessment only	90	9	64	27	86	6	60	34
	95% CI			(4-17)	(54-74)	(18-37)		(2-13)	(49-71)	(24-45)
Beutel et al. 2014	Treatment	Other - Psychodynamic therapy	59	7	22	71	Follow-up data not reported			
	95% CI			(2-16)	(12-35)	(58-82)				
	Control	Treatment as usual - Offered referral to GP for psychological or pharmacological treatment	58	14	47	40				
	95% CI			(6-25)	(33-60)	(27-53)				
Bredal et al. 2014	Treatment	Psychoeducation	169	15	54	30	174	9	47	44
	95% CI			(10-22)	(47-62)	(23-38)		(5-15)	(39-54)	(37-52)
	Control	Treatment as usual - Nurse-led support group	160	22	52	26	158	9	41	50
	95% CI			(16-29)	(44-60)	(20-34)		(5-15)	(33-49)	(42-58)

			Post treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
Carlson et al. 2013	Treatment	Mindfulness-based therapy	69	4	59	36	Follow-up data not reported			
	95% CI			(0-12)	(47-71)	(25-49)				
	Treatment	Supportive therapy	72	14	56	31				
	95% CI			(7-24)	(43-67)	(20-43)				
	Control	Stress management seminar	36	8	61	31				
	95% CI			(2-22)	(43-77)	(16-48)				
Classen et al. 2008	Treatment	Supportive therapy + education material	151	11	68	22	141	9	62	29
	95% CI			(6-17)	(59-75)	(16-29)		(4-14)	(54-70)	(22-37)
	Control	Education material	148	10	70	20	145	12	61	26
	95% CI			(6-16)	(62-77)	(14-27)		(8-19)	(53-69)	(19-34)
Desautels et al. 2018	Treatment	Cognitive behavioural therapy	25	0	24	76	27	4	19	78
	95% CI			(0-14)	(9-45)	(55-91)		(0-19)	(6-38)	(58-91)
	Control	Bright light therapy	22	0	32	68	25	0	40	60
	95% CI			(0-15)	(14-55)	(45-86)		(0-14)	(21-61)	(39-79)

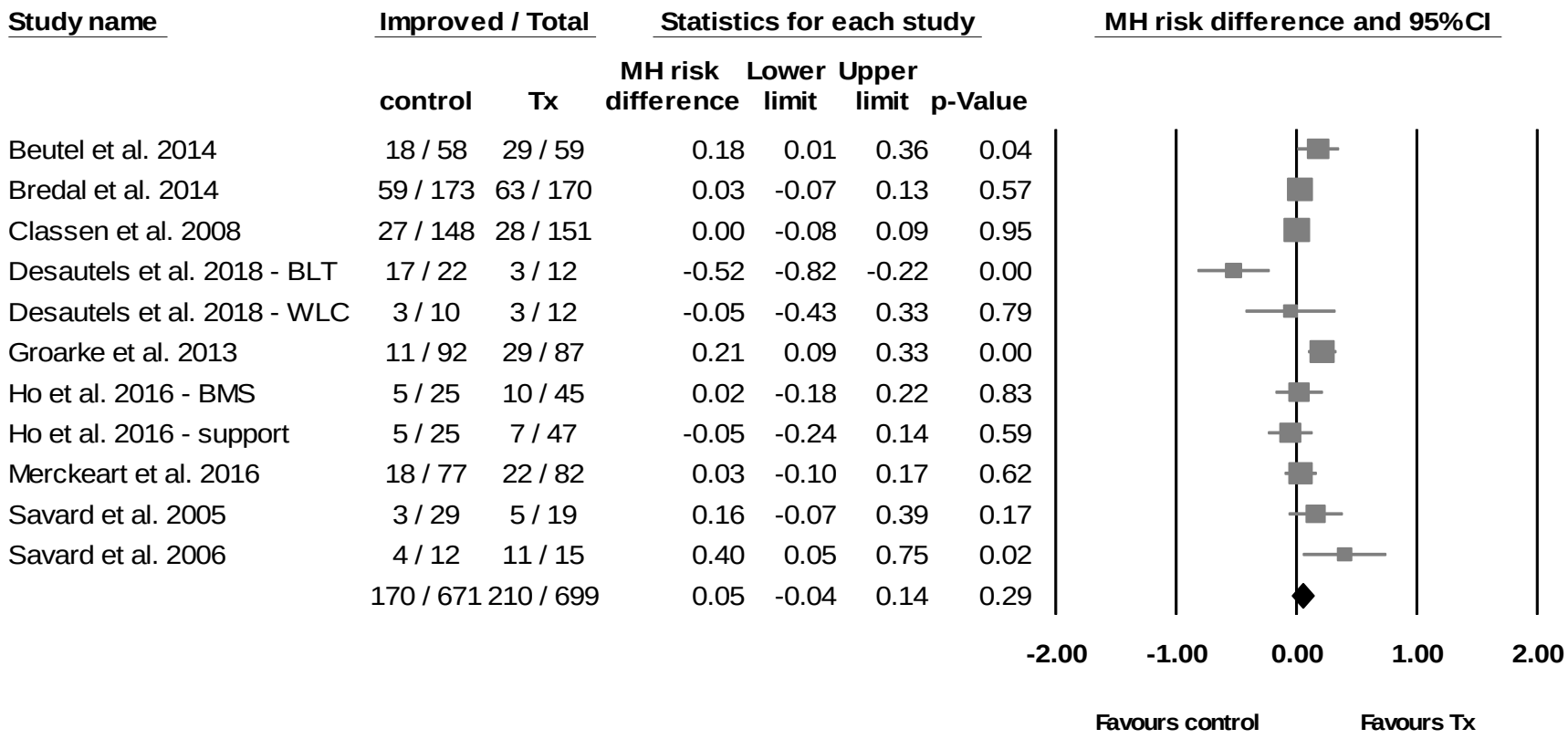
			Post treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
	Control	Wait-list control	10	10	50	40	Not applicable			
	95% CI			(0-45)	(19-81)	(12-74)				
Graves et al. 2003	Treatment	Cognitive behavioural therapy	7	29	29	43	Follow-up data not reported			
	95% CI			(4-71)	(4-71)	(10-82)				
	Control	Treatment as usual from medical community + educational material	6	17	83	0				
	95% CI			(0-64)	(36-100)	(0-46)				
Groarke et al. 2013	Treatment	Cognitive behavioural therapy	87	13	64	23	87	11	70	18
	95% CI			(6-21)	(53-74)	(15-33)		6-20%	59-79%	11-28%
	Control	Treatment as usual - Support from oncology nurses	92	17	73	10	92	22	62	16
	95% CI			(10-27)	(63-82)	(5-18)		14-32%	52-72%	9-25%
Ho et al. 2016	Treatment	Other - Body-mind-spirit intervention	45	11	73	16	47	11	74	15
	95% CI			(4-24)	(58-85)	(6-29)		4-23%	60-86%	6-28%
	Treatment	Supportive therapy	47	19	66	15	47	13	70	17

			Post treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
	95% CI			(9-33)	(51-79)	(6-28)		5-26%	55-83%	8-31%
	Control	Self-help support group	51	8	78	14	48	23	56	21
	95% CI			(2-19)	(65-89)	(6-26)		12-37%	41-71%	10-35%
Merckaert et al. 2016	Treatment	Cognitive behavioural therapy	82	7	63	29	Follow-up data not reported			
	95% CI			(3-15)	(52-74)	(20-40)				
	Control	Treatment as usual - Peer support group	77	9	71	19				
	95% CI			(4-18)	(60-81)	(11-30)				
Sandgren et al 2003; 2007	Treatment	Psychoeducation	81	15	60	25	77	65	27	8
	95% CI			(8-24)	(49-71)	(16-36)		53-75%	18-39%	3-16%
	Treatment	Other - Emotional expression	90	12	64	23	92	61	34	5
	95% CI			(6-21)	(54-74)	(15-33)		50-71%	24-44%	2-12%
	Control	Treatment as usual - Nurse help line was available	53	13	55	32	49	61	33	6
	95% CI			(5-25)	(40-68)	(20-46)		46-75%	20-48%	1-17%

			Post treatment (%)				Follow-up (%)			
Study name	**Treatment or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Total**	**Deteriorated**	**No change**	**Improved**
Savard et al. 2005	Treatment	Cognitive behavioural therapy	19	0	79	21	Not applicable			
	95% CI			(0-18)	(54-94)	(6-46)				
	Control	Wait-list control	29	17	72	10				
	95% CI			(6-36)	(53-87)	(2-27)				
Savard et al. 2006	Treatment	Cognitive behavioural therapy	15	0	20	80	Not applicable			
	95% CI			(0-22)	(4-48)	(52-96)				
	Control	Wait-list control	12	0	58	42				
	95% CI			(0-26)	(28-85)	(15-72)				

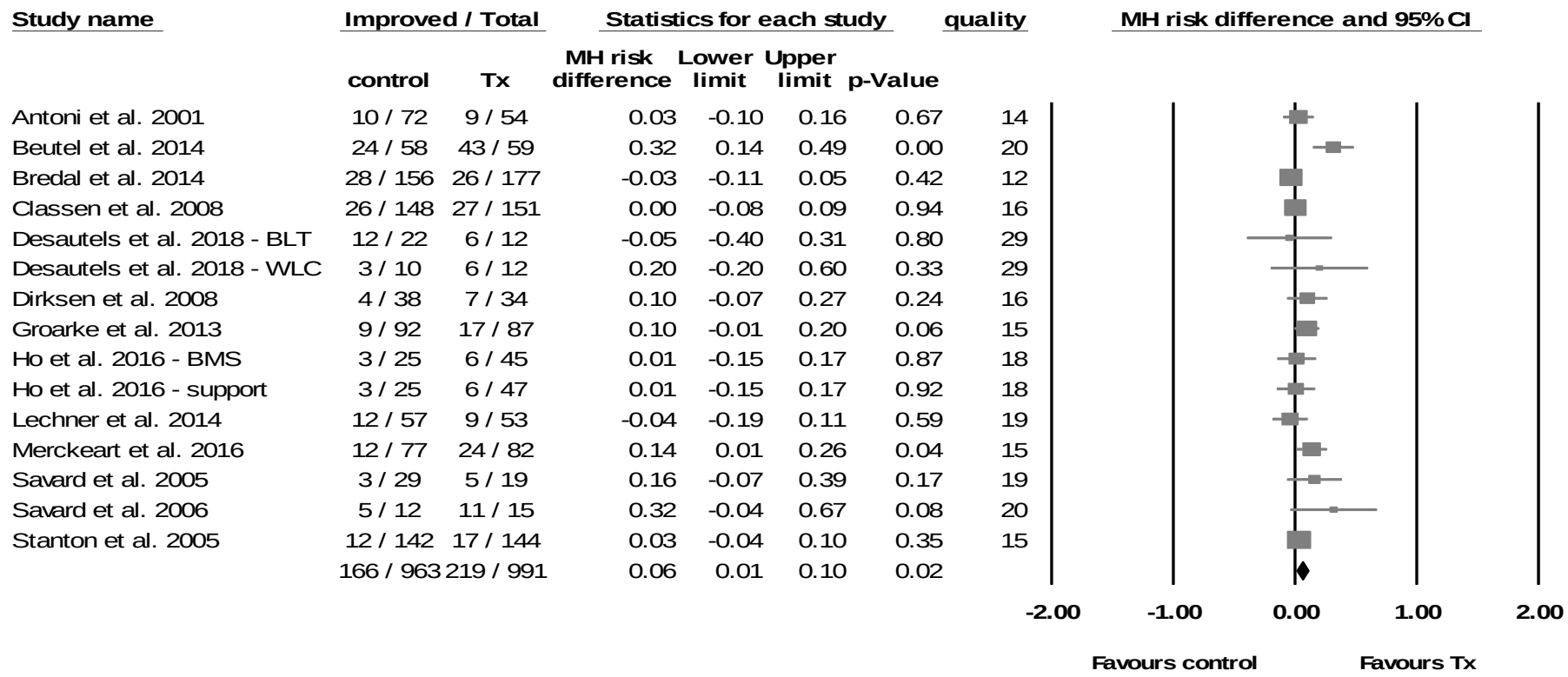
Note. n = number of patients; RCI = reliable change index at P<0.05; CI = confidence interval

Appendix 11:
Treatment effects at post-treatment and follow-up for anxiety, depression and general distress in the total sample based on risk differences for improvement (study 3)



Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy

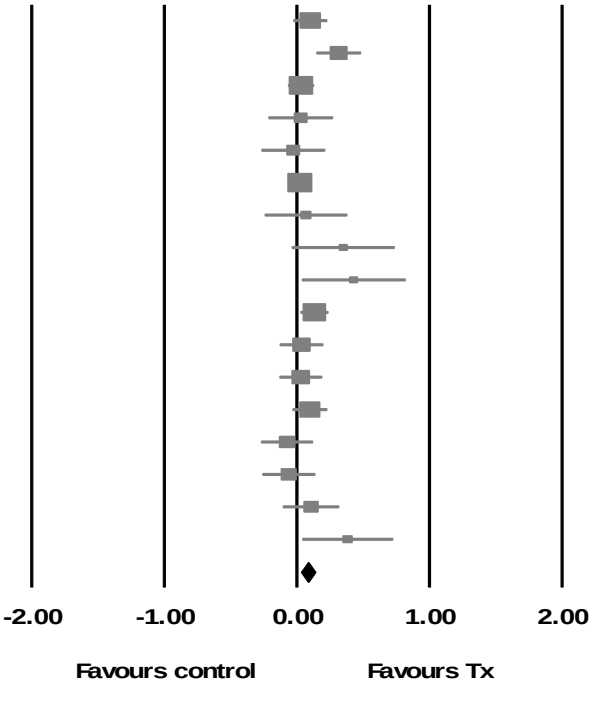
Figure 11A: Forest plot of risk differences for improvement for anxiety at post treatment in the total sample



Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy

Figure 11B: Forest plot of risk differences for improvement for depression at post treatment in the total sample

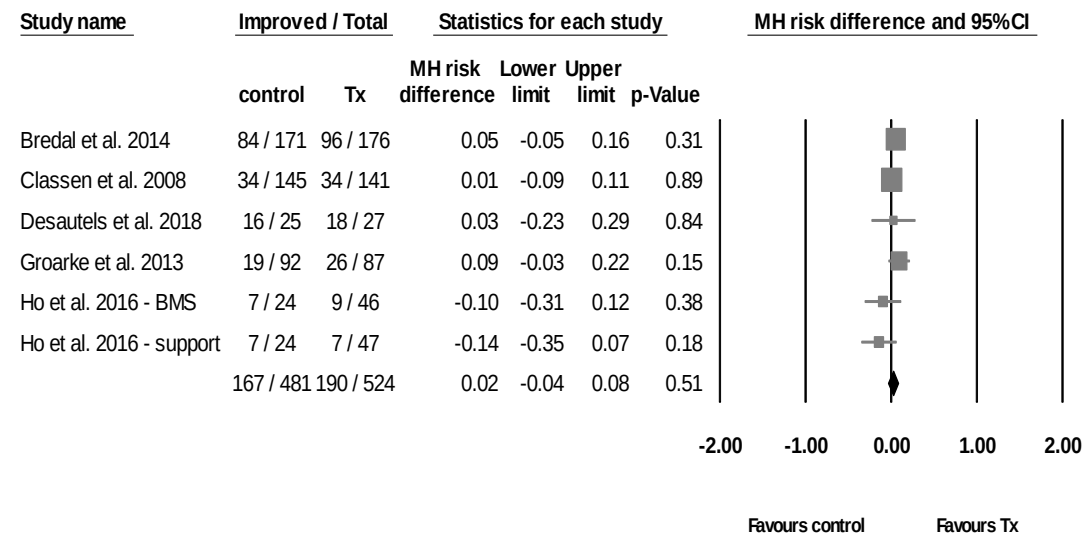
Study name	Improved / Total		Statistics for each study			
	control	Tx	MH risk difference	Lower limit	Upper limit	p-Value
Andersen et al. 2004	24 / 90	39 / 106	0.10	-0.03	0.23	0.13
Beutel et al. 2014	23 / 58	42 / 59	0.32	0.14	0.49	0.00
Bredal et al. 2014	42 / 153	51 / 167	0.03	-0.07	0.13	0.54
Carlson et al. 2013 - MBT	6 / 18	25 / 69	0.03	-0.22	0.27	0.82
Carlson et al. 2013 - support	6 / 18	22 / 72	-0.03	-0.27	0.21	0.82
Classen et al. 2008	29 / 148	33 / 151	0.02	-0.07	0.11	0.63
Desautels et al. 2018 - BLT	15 / 22	9 / 12	0.07	-0.24	0.38	0.67
Desautels et al. 2018 - WLC	4 / 10	9 / 12	0.35	-0.04	0.74	0.08
Graves et al. 2003	0 / 6	3 / 7	0.43	0.04	0.82	0.03
Groarke et al. 2013	9 / 92	20 / 87	0.13	0.02	0.24	0.02
Ho et al. 2016 - BMS	3 / 25	7 / 45	0.04	-0.13	0.20	0.67
Ho et al. 2016 - support	3 / 25	7 / 47	0.03	-0.13	0.19	0.73
Merckeart et al. 2016	15 / 77	24 / 82	0.10	-0.03	0.23	0.15
Sandgren et al. 2003 - EE	8 / 26	21 / 90	-0.07	-0.27	0.12	0.46
Sandgren et al. 2003 - PsyEd	8 / 26	20 / 81	-0.06	-0.26	0.14	0.55
Savard et al. 2005	3 / 29	4 / 19	0.11	-0.11	0.32	0.33
Savard et al. 2006	5 / 12	12 / 15	0.38	0.04	0.73	0.03
	203 / 835	348 / 1121	0.08	0.03	0.13	0.00



Note.

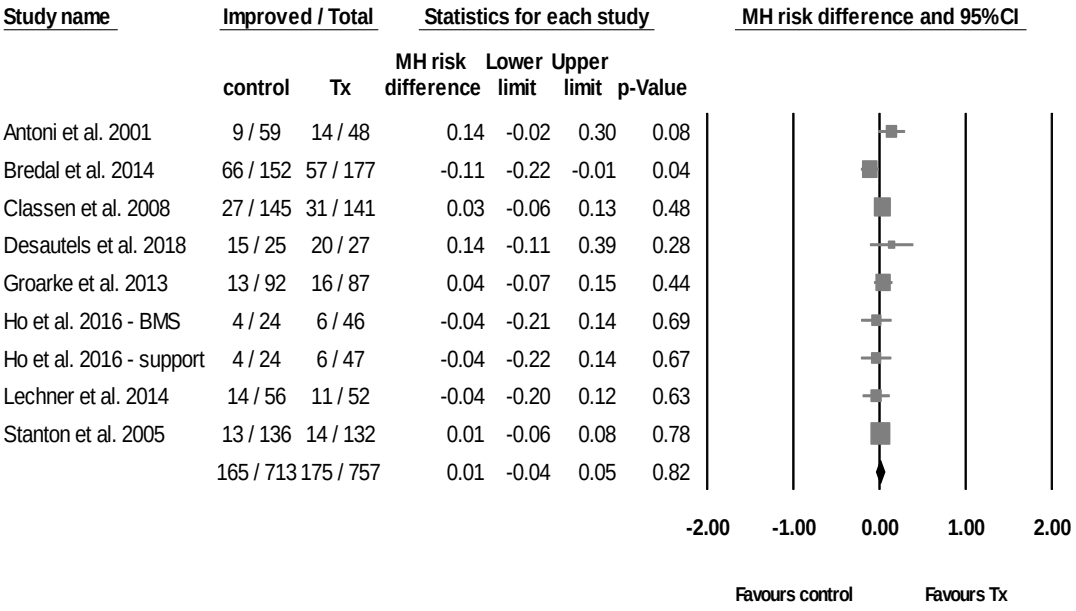
BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy; EE = emotional expression; PsyEd = psychoeducation

Figure 11C: Forest plot of risk differences for improvement for general distress at post treatment in the total sample



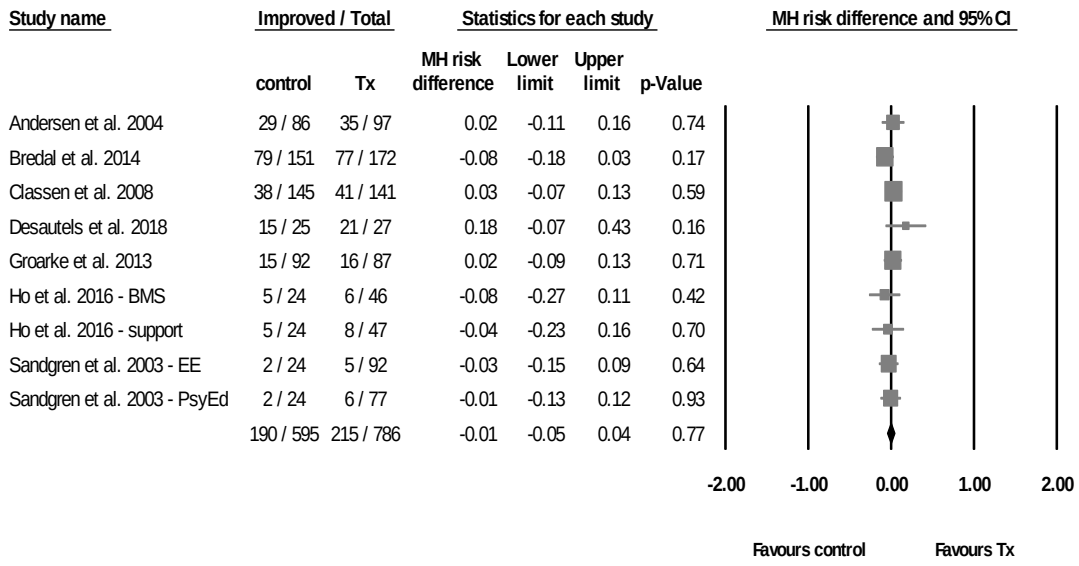
Note. BMS = body mind spirit; support = supportive therapy

Figure 11D. Forest plot of risk differences for improvement for anxiety at follow-up in the total sample



Note. BMS = body mind spirit; support = supportive therapy

Figure 11E: Forest plot of risk differences for improvement for depression at follow-up in the total sample



Note. BMS = body mind spirit; support = supportive therapy; EE = emotional expression; PsyEd = psychoeducation

Figure 11F: Forest plot of risk differences for improvement for general distress at follow-up in the total sample

Appendix 12:

Classification of patients at post-treatment and follow-up according to Jacobson's first and second (RCI) criteria for anxiety, depression and general distress in the distressed sub-sample (study 3)

Table 12A: Classification of patients at post-treatment and follow-up according to Jacobson's first and second (RCI) criteria for anxiety in the distressed sub-sample

			Post-treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
Beutel et al. 2014	Tx	Other - Psychodynamic therapy	54	9	39	28	24	Follow-up data not reported				
	95% CI			(3-20)	(26-53)	(16-42)	(13-38)					
	Control	Treatment as usual - Offered referral to GP for psychological or pharmacological treatment	50	8	58	26	8					
	95% CI			(2-19)	(43-72)	(15-40)	(2-19)					
Bredal et al. 2014	Tx	Psychoeducation	90	12	40	16	32	90	4	22	21	52
	95% CI			(6-21)	(30-51)	(9-25)	(23-43)		(1-11)	(14-32)	(13-31)	(41-63)
	Control	Treatment as usual - Nurse-led support group	106	13	47	11	28	101	5	28	14	53
	95% CI			(7-21)	(37-57)	(6-19)	(20-38)		(2-11)	(19-38)	(8-22)	(43-63)
Classen et al. 2008	Tx	Supportive therapy + education material	56	13	43	18	27	49	12	43	27	18
	95% CI			(5-24)	(30-57)	(9-30)	(16-40)		(5-25)	(29-58)	(15-41)	(9-32)

			Post-treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
	Control	Education material	57	9	60	14	18	55	7	51	16	25
	95% CI			(3-19)	(46-72)	(6-26)	(19-30)		(2-18)	(37-65)	(8-29)	(15-39)
Desautels et al. 2018	Tx	Cognitive behavioural therapy	19	0	68	16	16	7	0	57	0	43
	95% CI			(0-18)	(43-87)	(3-40)	(3-40)		(0-41)	(18-90)	(0-41)	(10-82)
	Control	Bright light therapy	17	6	0	41	53	5	0	80	0	20
	95% CI			(0-29)	(0-20)	(18-67)	(28-77)		(0-52)	(28-99)	(0-52)	(0-72)
	Control	Wait-list therapy	6	0	50	50	0		Not applicable			
	95% CI			(0-46)	(12-88)	(12-88)	(0-46)					
Groarke et al. 2013	Tx	Cognitive behavioural therapy	35	0	34	11	54	35	17	23	20	40
	95% CI			(0-10)	(19-52)	(3-27)	(37-71)		(7-34)	(10-40)	(8-37)	(24-58)
	Control	Treatment as usual - Support from oncology nurses	28	7	75	7	11	28	7	46	11	36
	95% CI			(0-24)	(55-89)	(0-24)	(2-28)		(0-24)	(28-66)	(2-28)	(19-56)
Ho et al. 2016	Tx	Other - Body-mind-spirit intervention	16	6	56	6	31	15	20	53	7	20

			Post-treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
	95% CI			(0-30)	(30-80)	(0-30)	(11-59)		(4-48)	(27-79)	(0-32)	(4-48)
	Tx	Supportive therapy	22	23	55	5	18	22	9	68	9	14
	95% CI			(8-45)	(32-76)	(0-23)	(5-40)		(1-29)	(45-86)	(1-29)	(3-35)
	Control	Self-help support group	23	13	52	13	22	22	23	36	18	23
	95% CI			(3-34)	(31-73)	(3-34)	(7-44)		(8-45)	(17-59)	(5-40)	(8-45)
Merckaert et al. 2016	Tx	Cognitive behavioural therapy	40	5	50	15	30	Follow-up data not reported				
	95% CI			(0-17)	(34-66)	(6-30)	(17-47)					
	Control	Treatment as usual - peer support group	43	5	63	14	19					
	95% CI			(0-16)	(47-77)	(5-28)	(8-33)					
Savard et al. 2005	Tx	Cognitive behavioural therapy	9	0	44	11	44	Not applicable				
	95% CI			(0-34)	(14-79)	(0-48)	(14-79)					
	Control	Wait-list control	12	33	50	0	17					
	95% CI			(10-65)	(21-79)	(0-26)	(2-48)					

			Post-treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
Savard et al. 2006	Tx	Cognitive behavioural therapy	12	0	8	42	50	Not applicable				
	95% CI			(0-26)	(0-38)	(15-72)	(21-79)					
	Control	Wait-list control	11	18	45	27	9					
	95% CI			(2-52)	(17-77)	(6-61)	(0-41)					

Note. Tx = treatment; n = number of patients; RCI = reliable change index at P<0.05; CI = confidence interval

Table 12B: Classification of patients at post-treatment and follow-up according to Jacobson's first and second (RCI) criteria for depression in the distressed sub-sample

			Post treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
Antoni et al. 2001	Tx	Cognitive behavioural therapy	18	11	44	17	28	17	0	35	29	35
	95% CI			(1-35)	(22-69)	(4-41)	(10-53)		(0-20)	(14-62)	(10-56)	(14-62)
	Control	Condensed version of active treatment	22	0	59	23	18	18	6	50	33	11
	95% CI			(0-15)	(36-79)	(8-45)	(5-40)		(0-27)	(26-74)	(13-59)	(1-35)
Beutel et al. 2014	Tx	Other - Psychodynamic therapy	59	3	24	37	36	Follow-up data not reported				
	95% CI			(0-12)	(14-37)	(25-51)	(24-49)					
	Control	Treatment as usual - Offered referral to GP for psychological or pharmacological treatment	58	17	41	21	21					
	95% CI			(9-29)	(29-55)	(11-33)	(11-33)					
Bredal et al. 2014	Tx	Psychoeducation	47	4	62	6	28	45	2	27	0	71

			Post treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
	95% CI			(0-15)	(46-75)	(1-18)	(16-43)		(0-12)	(15-42)	(0-8)	(56-84)
	Control	Treatment as usual - Nurse-led support group	52	4	63	8	25	47	2	13	9	77
	95% CI			(0-13)	(49-76)	(2-19)	(14-39)		(0-11)	(5-26)	(2-20)	(62-88)
Classen et al. 2008	Tx	Supportive therapy + education material	26	4	42	19	35	28	0	46	21	32
	95% CI			(0-20)	(23-63)	(7-39)	(17-56)		(0-12)	(28-66)	(8-41)	(16-52)
	Control	Education material	18	11	28	17	44	19	11	47	5	37
	95% CI			(1-35)	(10-53)	(4-41)	(22-69)		(1-33)	(24-71)	(0-26)	(16-62)
Desautels et al. 2018	Tx	Cognitive behavioural therapy	19	0	42	16	42	19	0	21	11	68
	95% CI			(0-18)	(20-67)	(3-40)	(20-67)		(0-18)	(6-46)	(1-33)	(43-87)
	Control	Bright light therapy	15	0	27	33	40	16	0	19	25	56
	95% CI			(0-22)	(8-55)	(12-62)	(16-68)		(0-21)	(4-46)	(7-52)	(30-80)
	Control	Wait-list control	8	0	63	13	25	Not applicable				
	95% CI			(0-37)	(24-91)	(0-53)	(3-65)					

			Post treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
Dirksen et al. 2008	Tx	Cognitive behavioural therapy	13	0	54	15	31	Follow-up data not reported				
	95% CI			(0-25)	(25-81)	(2-45)	(9-61)					
	Control	Educational components of active treatment	5	20	40	0	40					
	95% CI			(0-72)	(5-85)	(0-52)	(5-85)					
Groarke et al. 2013	Tx	Cognitive behavioural therapy	12	0	67	0	33	12	17	42	8	33
	95% CI			(0-26)	(35-90)	(0-26)	(10-65)		(2-48)	(15-72)	(0-38)	(10-65)
	Control	Treatment as usual - Support from oncology nurses	12	0	58	0	42	12	8	25	33	33
	95% CI			(0-26)	(28-85)	(0-26)	(15-72)		(0-38)	(5-57)	(10-65)	(10-65)
Ho et al. 2016	Tx	Other - Body-mind-spirit intervention	5	0	40	20	40	5	0	20	20	60
	95% CI			(0-52)	(5-85)	(0-72)	(5-85)		(0-52)	(0-72)	(0-72)	(15-95)
	Tx	Supportive therapy	9	33	44	0	22	9	0	78	0	22
	95% CI			(7-70)	(14-79)	(0-34)	(3-60)		(0-34)	(40-97)	(0-34)	(3-60)

			Post treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
	Control	Self-help support group	7	14	71	14	0	7	29	43	14	14
	95% CI			(0-58)	(29-96)	(0-58)	(0-41)		(4-71)	(10-82)	(0-58)	(0-58)
Lechner et al. 2016	Tx	Cognitive behavioural therapy	53	8	75	17	0	52	4	75	21	0
	95% CI			(2-18)	(62-86)	(8-30)	(0-7)		(0-13)	(61-86)	(11-35)	(0-7)
	Control	Educational information delivered by therapist	57	12	67	21	0	56	9	66	25	0
	95% CI			(5-24)	(53-79)	(11-34)	(0-6)		(3-20)	(52-78)	(14-38)	(0-6)
Merckaert et al. 2016	Tx	Cognitive behavioural therapy	21	0	33	19	48	Follow-up data not reported				
	95% CI			0-16%	15-57%	5-42%	26-70%					
	Control	Treatment as usual - Peer support group	15	0	53	7	40					
	95% CI			(0-22)	(27-79)	(0-32)	(16-68)					
Savard et al. 2005	Tx	Cognitive behavioural therapy	2	0	0	0	100	Not applicable				
	95% CI			(0-84)	(0-84)	(0-84)	(16-100)					

			Post treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
	Control	Wait-list control	3	0	33	0	67					
	95% CI			(0-71)	(0-91)	(0-71)	(9-99)					
Savard et al. 2006	Tx	Cognitive behavioural therapy	10	0	10	30	60	Not applicable				
	95% CI			(0-31)	(0-45)	(7-65)	(26-88)					
	Control	Wait-list control	9	0	56	0	44					
	95% CI			(0-34)	(21-86)	(0-34)	(14-79)					
Stanton et al. 2005	Tx	Psychoeducation + educational material	40	8	53	25	15	38	5	63	26	5
	95% CI			(2-20)	(36-68)	(13-41)	(6-30)		(0-18)	(46-78)	(13-43)	(0-18)
	Control	Education material	24	4	67	21	0	24	4	50	8	38
	95% CI			(0-21)	(53-79)	(11-34)	(0-6)		(0-21)	(29-71)	(1-27)	(19-59)

Note. Tx = treatment; n = number of patients; RCI = reliable change index at P<0.05; CI = confidence interval

Table 12C: Classification of patients at post-treatment and follow-up according to Jacobson's first and second (RCI) criteria for general distress in the distressed sub-sample

			Post treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
Andersen et al. 04	Tx	Other - Biobehavioral intervention	48	13	27	40	21	44	2	41	25	32
	95% CI			(5-25)	(15-42)	(26-55)	(10-35)		(0-12)	(26-57)	(13-40)	(19-48)
	Control	Assessment only	30	7	37	47	10	29	3	34	38	24
	95% CI			(0-22)	(20-56)	(28-66)	(2-27)		(0-18)	(18-54)	(21-58)	(10-44)
Beutel et al. 2014	Tx	Other - Psychodynamic therapy	59	7	22	51	20	Follow-up data not reported				
	95% CI			(2-16)	(12-35)	(37-64)	(11-33)					
	Control	Treatment as usual - Offered referral to GP for psychological or pharmacological treatment	57	14	47	32	7					
	95% CI			(6-26)	(34-61)	(20-45)	(2-17)					
Bredal et al. 2014	Tx	Psychoeducation	81	7	51	15	27	81	4	22	21	53
	95% CI			(3-15)	(39-62)	(8-24)	(18-38)		(0-10)	(14-33)	(13-31)	(42-64)

			Post treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
	Control	Treatment as usual - Nurse-led support group	80	15	50	8	28	78	4	15	26	55
	95% CI			(8-25)	(39-61)	(3-16)	(18-39)		(0-11)	(8-25)	(16-37)	(43-66)
Carlson et al. 2013	Tx	Mindfulness-based therapy	27	7	26	19	48	Follow-up data not reported				
	95% CI			(0-24)	(11-46)	(6-38)	(29-68)					
	Tx	Supportive therapy	29	17	34	28	21					
	95% CI			(6-36)	(18-54)	(13-47)	(8-40)					
	Control	Stress management seminar	14	14	29	21	36					
	95% CI			(2-43)	(8-58)	(5-51)	(13-65)					
Classen et al. 2008	Tx	Supportive therapy + education material	49	12	37	20	31	45	4	33	33	29
	95% CI			(5-25)	(23-52)	(10-34)	(18-45)		(0-15)	(20-49)	(20-49)	(16-44)
	Control	Education material	44	5	45	23	27	42	17	26	26	31
	95% CI			(0-15)	(30-61)	(11-38)	(15-43)		(7-31)	(14-42)	(14-42)	(18-47)
Desautels et al. 2018	Tx	Cognitive behavioural therapy	23	0	17	43	39	24	0	17	33	50

			Post treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
	95% CI			(0-15)	(5-39)	(23-66)	(20-61)		(0-14)	(5-37)	(16-55)	(29-71)
	Control	Bright light therapy	21	0	29	33	38	24	0	38	25	38
	95% CI			(0-16)	(11-52)	(15-57)	(18-62)		(0-14)	(19-59)	(10-47)	(19-59)
	Control	Wait-list control	9	11	56	33	0	Not applicable				
	95% CI			(0-48)	(21-86)	(7-70)	(0-34)					
Graves et al. 2003	Tx	Cognitive behavioural therapy	3	33	33	33	0	Follow-up data not reported				
	95% CI			(0-91)	(0-91)	(0-91)	(0-71)					
	Control	Treatment as usual from medical community + educational material	0	0	0	0	0					
	95% CI		0	0	0	0	0					
Groarke et al. 2013	Tx	Cognitive behavioural therapy	28	0	0.5	0.14	0.36	28	4	50	11	36
	95% CI			(0-12)	(31-69)	(4-33)	(19-56)		(0-18)	(31-69)	(2-28)	(19-56)
	Control	Treatment as usual - Support from oncology nurses	25	8	80	4	8	25	12	48	24	16
	95% CI			(0-26)	(59-93)	(0-20)	(0-26)		(3-31)	(28-69)	(9-45)	(5-36)

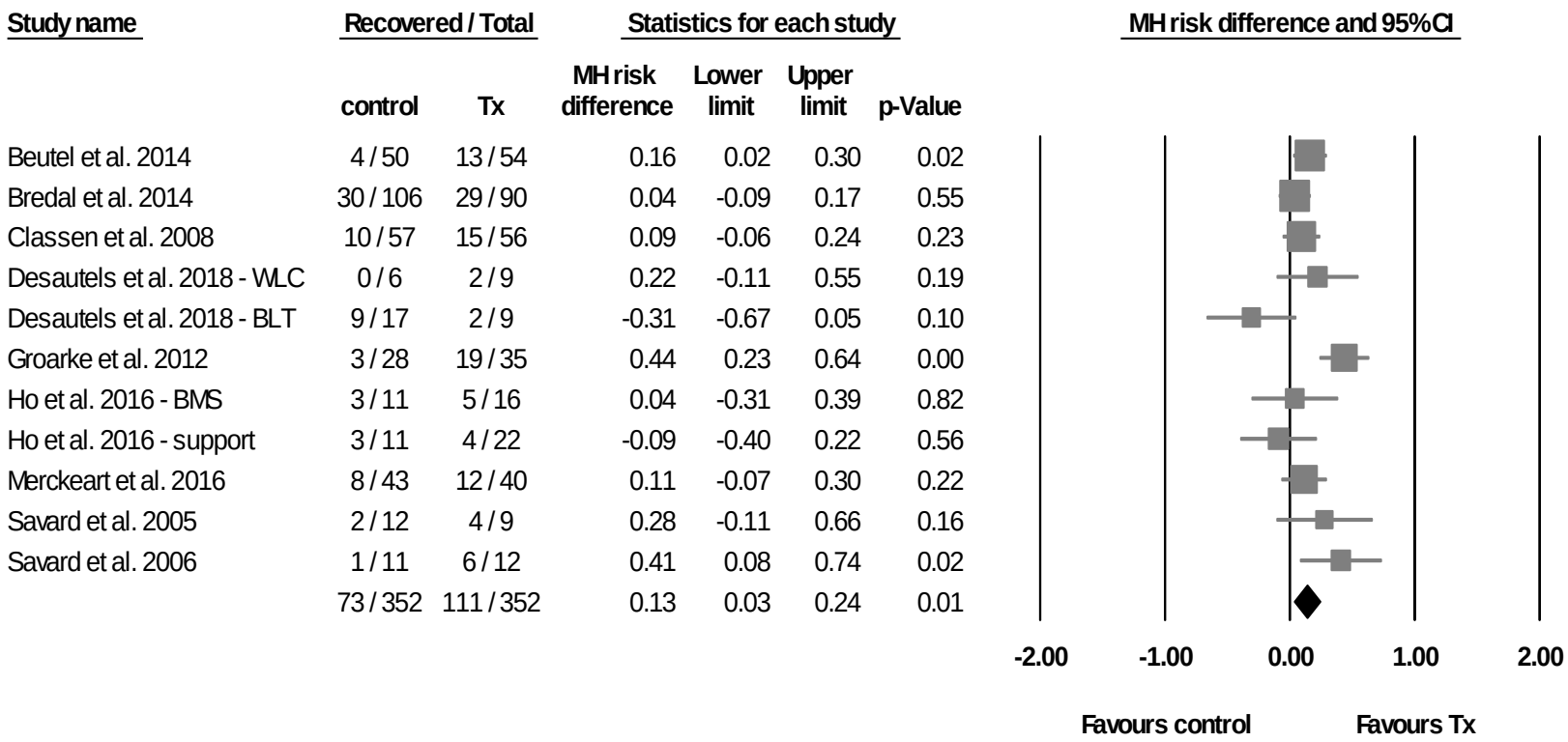
			Post treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
Ho et al. 2016	Tx	Other - body-mind-spirit intervention	14	0	71	14	14	13	0	69	23	8
	95% CI			(0-23)	(42-92)	(2-43)	(2-43)		(0-25)	(39-91)	(5-54)	(0-36)
	Tx	Supportive therapy	24	21	54	4	21	24	13	54	8	25
	95% CI			(7-42)	(33-74)	(0-21)	(7-42)		(3-32)	(33-74)	(1-27)	(10-47)
	Control	Self-help support group	20	5	75	15	5	19	26	37	32	5
	95% CI			(0-25)	(51-91)	(3-38)	(0-25)		(9-51)	(16-62)	(13-57)	(0-26)
Merckaert et al. 2016	Tx	Cognitive behavioural therapy	39	8	38	28	26	Follow-up data not reported				
	95% CI			(2-21)	(23-55)	(15-45)	(13-42)					
	Control	Treatment as usual - peer support group	32	13	53	22	13					
	95% CI			(4-29)	(35-71)	(9-40)	(4-29)					
Sandgren et al 2003; 2007	Tx	Psychoeducation	22	9	32	14	45	21	19	52	29	0
	95% CI			(1-29)	(14-55)	(3-35)	(24-68)		(5-42)	(30-74)	(11-52)	(0-16)
	Tx	Other - Emotional expression	30	7	43	23	27	30	20	63	17	0

			Post treatment (%)					Follow-up (%)				
Study name	**Tx or control**	**Type**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**	**Total**	**Deteriorated**	**No change**	**Improved**	**Recovered**
	95% CI			(0-22)	(25-63)	(10-42)	(12-46)		(8-39)	(44-80)	(6-35)	(0-12)
	Control	Treatment as usual - Nurse help line was available	21	5	29	43	24	20	25	60	15	0
	95% CI			(0-24)	(11-52)	(22-66)	(8-47)		(9-49)	(36-81)	(3-38)	(0-17)
Savard et al. 2005	Tx	Cognitive behavioural therapy	8	0	63	13	25	Not applicable				
	95% CI			(0-37)	(24-91)	(0-53)	(3-65)					
	Control	Wait-list control	7	14	71	14	0					
	95% CI			(0-58)	(29-96)	(0-58)	(0-41)					
Savard et al. 2006	Tx	Cognitive behavioural therapy	13	0	8	54	38	Not applicable				
	95% CI			(0-25)	(0-36)	(25-81)	(14-68)					
	Control	Wait-list control	11	0	64	27	9					
	95% CI			(0-28)	(31-89)	(6-61)	(0-41)					

Note. Tx = treatment; n = number of patients; RCI = reliable change index at P<0.05; CI = confidence interval

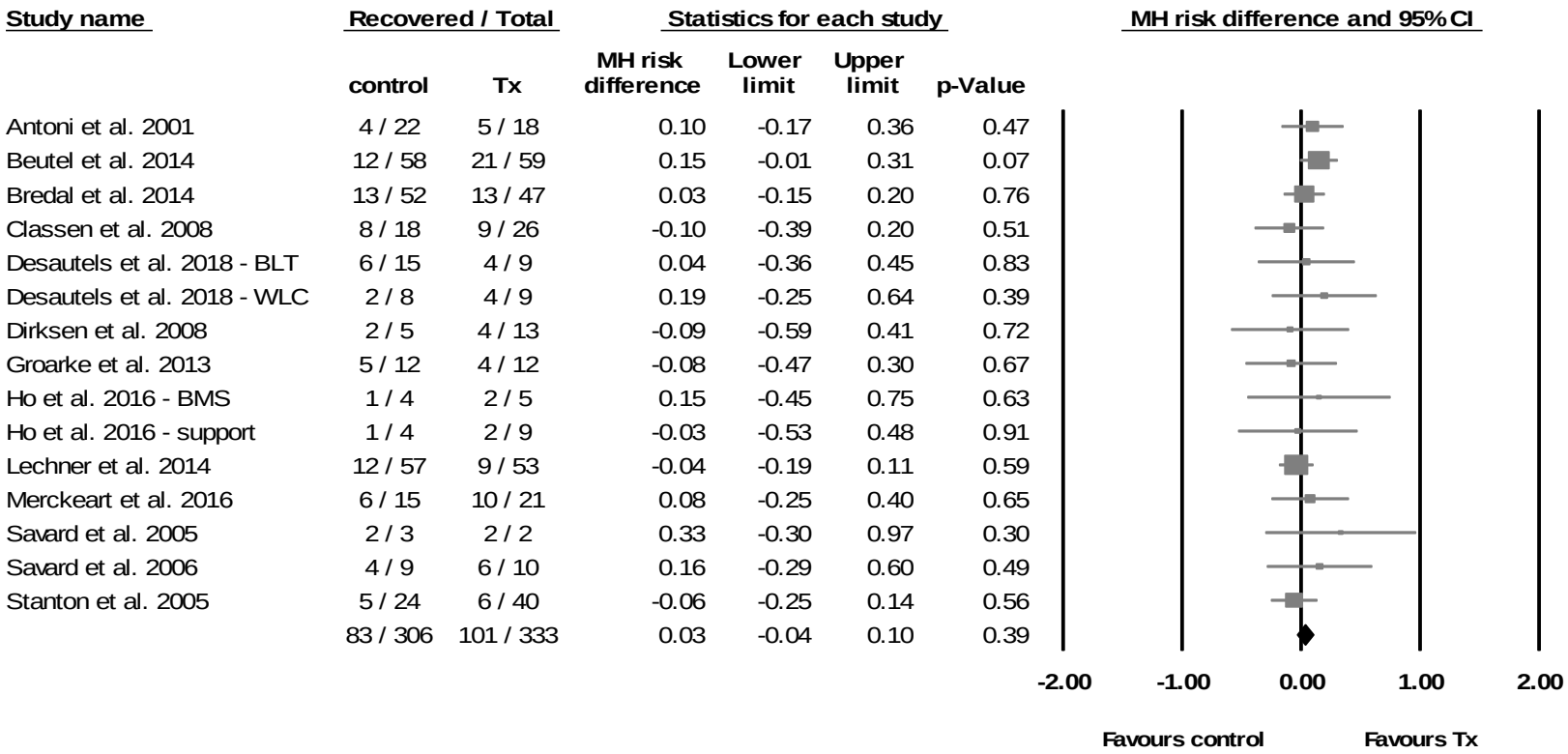
Appendix 13:

Treatment effects at post-treatment and follow-up for anxiety, depression and general distress in the distressed sub-sample based on risk differences for recovery (study 3)



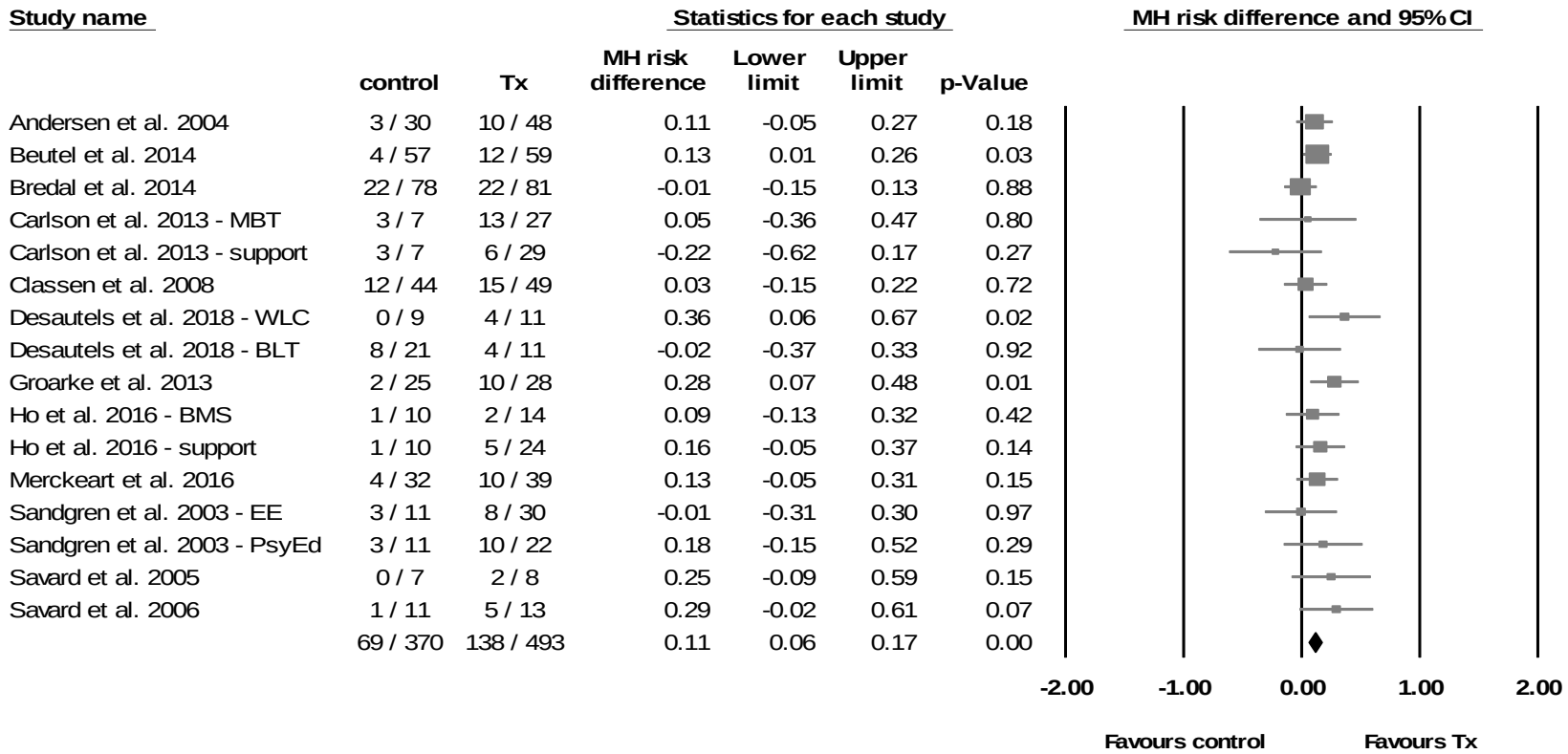
Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy

Figure 13A: Forest plot of risk differences for recovery for anxiety at post treatment in the distressed sub-sample



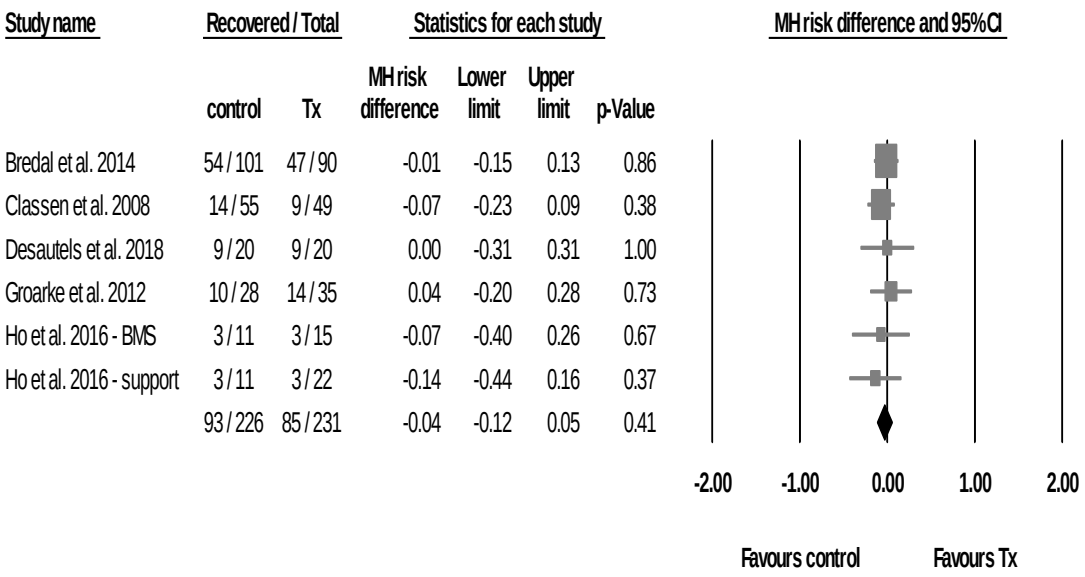
Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy

Figure 13B: Forest plot of risk differences for recovery for depression at post treatment in the distressed sub-sample



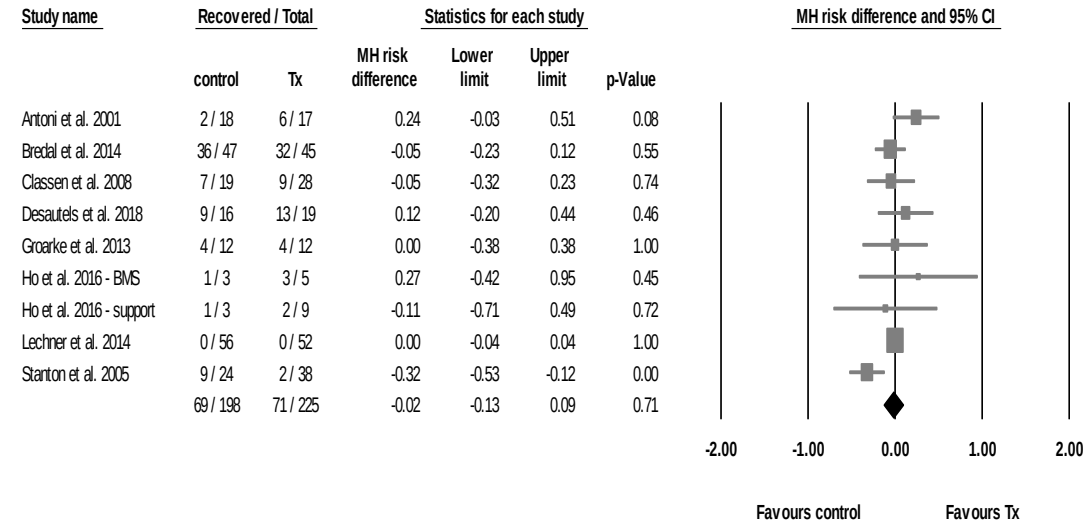
Note. BLT = bright light therapy; MBT = mindfulness-based therapy; WLC = waitlist control; BMS = body mind spirit; support = supportive therapy; EE = emotional expression; PsyEd = psychoeducation

Figure 13C: Forest plot of risk differences for recovery for general distress at post treatment in the distressed sub-sample



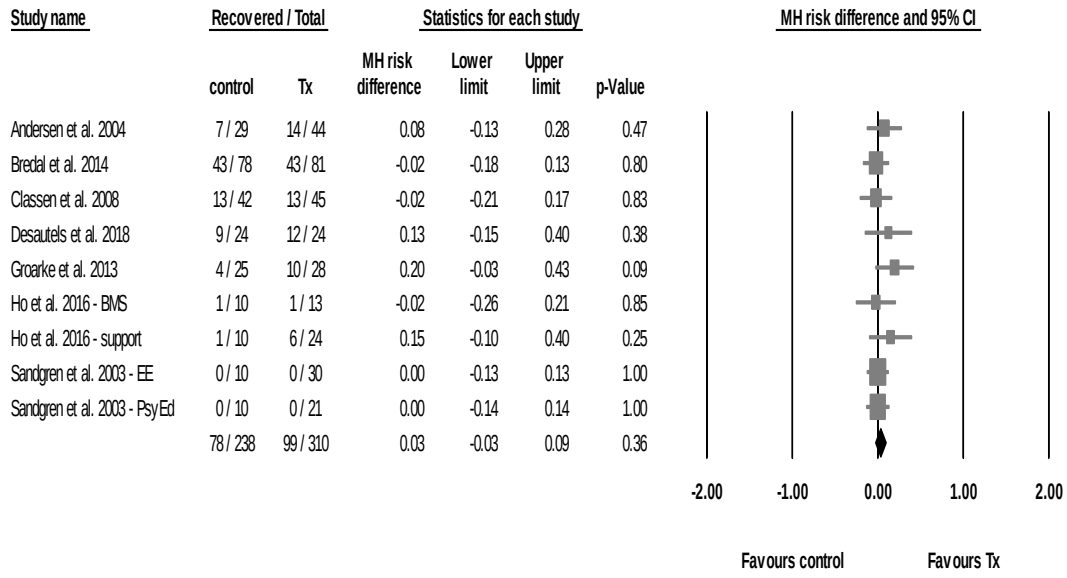
Note. BLT = BMS = body mind spirit; support = supportive therapy

Figure 13D: Forest plot of risk differences for recovery for anxiety at follow-up in the distressed sub sample



Note. BLT = BMS = body mind spirit; support = supportive therapy

Figure 13E: Forest plot of risk differences for recovery for depression at follow-up in the distressed sub-sample



Note. BLT = BMS = body mind spirit; support = supportive therapy; EE = emotional expression; PsyEd = psychoeducation

Figure 13F: Forest plot of risk differences for recovery for general distress at follow-up in the distressed sub-sample