Understanding the Impact of Baseline Patient Characteristics on Outcomes following Psychological Treatment for Depression.

By: Margaret Lyons (Klimyuk)

A thesis submitted in partial fulfilment of the requirements for the degree of Doctorate in Clinical Psychology

Clinical Psychology Unit, Department of Psychology

The University of Sheffield

May 2023

Declaration

I, the author, confirm this thesis is my own work and that I am aware of the University of Sheffield guidance on unfair means (www.sheffield.ac.uk/new-students/unfair-means). This thesis is submitted in partial fulfilment of the requirements for the degree of Doctorate in Clinical Psychology. This work has not been submitted for any other degree or to any other institution.

Structure and Word Count

Section One: Literature Review

Excluding references and tables:	7,999
Including references and tables:	14,656

Section Two: Empirical Study

Excluding references and tables:	7,661
Including references and tables:	11,010

Total

Excluding references and tables:	15,660
Including references and tables:	25,666

Lay Summary

Nearly a quarter of people in the United Kingdom are estimated to experience depression over the course of their lifetime. Depression is recognised to have a significant impact on a person's quality of life, but also on those around the person and wider society. Several forms of psychotherapy are recommended for the treatment of depression. Although these different forms of psychotherapy are similar in effectiveness, some people may respond better to one treatment over another. However, there is limited guidance on how to decide which form of psychotherapy should be offered to a patient. Additionally, many people do not improve following treatment and their depression continues. To increase successful treatment rates, researchers are exploring how baseline patient characteristics (i.e., characteristics measured at the beginning of treatment), influence psychotherapy outcomes. It is hoped that treatment outcomes can be improved by matching the most effective treatment to patients, based on their baseline characteristics.

The first chapter explored which psychosocial characteristics influence how people with persisting forms of depression respond to psychological therapy. Persisting forms of depression refers to a group of patients who have experienced depression for at least two years, or who have tried treatment but shown no or minimal response. A search of the existing literature for published studies on this topic was completed. Twenty-three studies were found, which altogether examined a total of sixty-five different variables. Most variables, such as sociodemographic factors (e.g., gender) were not found to influence psychotherapy outcomes. Only depression severity was repeatedly associated with psychotherapy outcomes, whereby higher depression severity at the beginning of treatment was found predictive of poorer outcomes. However, findings need to be replicated by further studies before they can be used to inform treatment allocation in clinical practice.

The second chapter aimed to understand whether complex cases would respond better to cognitive behavioural therapy or counselling for depression. These two psychotherapies are commonly offered to patients who access support for depression within the National Health Service (NHS) Talking Therapies services. The term ‘complex cases’ describes patients who present with several psychosocial characteristics that are associated with poorer outcomes. Pre-existing data from a trial completed in NHS Talking Therapies services was retrospectively analysed. This analysis found that complex cases responded similarly to cognitive behavioural therapy and counselling for depression. Additionally, over half of complex cases were found to show meaningful benefits from accessing psychotherapy for depression. Findings suggest that for complex cases depression severity can indeed improve following psychotherapy, and that either cognitive behavioural therapy or counselling for depression can be tried.

Acknowledgements

Firstly, I would like to thank my research supervisor Professor Jaime Delgadillo whose guidance and knowledge has helped this thesis feel manageable, and even (dare I say!) enjoyable. I have valued working with you and have learned a lot.

Thank you to everyone who participated in the StratCare trial. Without you this project would not exist.

To my friends and family, thank you believing in me. To my incredible friends, Aditi, Charlotte, and Lucy - I am honoured that we got to share this journey together, thank you for your endless support!

To my husband David, who has been my biggest supporter on this journey towards becoming a qualified clinical psychologist. Thank you for your continuous encouragement and patience over the years, for always listening to me, and for reminding me of the joys in life no matter how stressed I was.

And finally, to the world's best spaniel Benji. Always by my side, always ready to interrupt me for those all-essential cuddles, and always managing to put a smile on my face.

List of Contents

Part One: Literature Review

Predictors and Moderators of Change following Psychological Interventions for Persisting Forms of Depression: A Systematic Review.

Abstract

Background: Predictors and moderators of treatment response have been extensively studied within the field of major depressive disorder, with the aim of better matching patients to treatments. However, there is less understanding of predictors and moderators of response to psychological treatment for persisting forms of depression, such as chronic, recurrent, and treatment-resistant depression.

Methods: A systematic review (PROSPERO registration number CRD42022379257) was conducted by searching Web of Science, Scopus and PsycInfo up to the 1st of December 2022. A total of 23 eligible studies were identified. The Critical Appraisal Skills Programme checklists were used for risk of bias assessments. Narrative synthesis was used to summarise findings on predictors and moderators of response to psychological treatment for adults with persisting forms of depression.

Results: Sixty-five predictor and moderators variables were examined across included studies, categorised into sociodemographic, clinical, interpersonal/ personality, psychological and treatment variables. Findings were non-significant for 57%, significant for 18% and inconclusive for 25% of examined variables. Most variables were only examined by single studies. Amongst variables studied more frequently (age, gender, baseline depression severity, childhood trauma), only baseline depression severity was found to be a promising predictor of outcomes. Risk of bias was low-to-medium for the majority of studies.

Conclusion: Understanding of significant predictors and moderators for persisting forms of depressions is limited. Clinicians should be cautious when allocating patients with persisting forms of depression to psychological treatments based on baseline characteristics.

Keywords*: Depression, Predictors, Moderators, Psychosocial Intervention, Treatment-Resistant, Chronic, Recurrent*

Practitioner Points

- Baseline depression severity is a possible predictor of outcomes for patients with persisting forms of depression.
- At present there is insufficient evidence to support treatment allocation of patients with persisting forms of depression based on baseline characteristics.
- Further research is required to establish robust predictors and moderators of treatment response.

Introduction

Major depressive disorder (MDD) is a common mental health problem worldwide. The lifetime prevalence of major depression is estimated to be 25.8% in the United Kingdom (UK), with a probable 19.4% prevalence of recurrent MDD (Smith et al., 2013). MDD is associated with significant functional impairment, increased morbidity, and high societal cost (McLaughlin, 2011).

Persisting Forms of Depression

For many, MDD persists or re-occurs. It is estimated that approximately 50% of individuals with MDD will experience a chronic or recurrent course, and 20% will experience treatment-resistant depression (TRD; Crown et al., 2002). Numerous terms are used within the literature to describe such persisting forms of depression, with three commonly discussed MDD subtypes described in the following sections.

Chronic Depression

The occurrence of major depressive symptoms for at least two years is commonly referred to as chronic depression. In the current version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) by the American Psychiatric Association (APA, 2013), chronic depression is now conceptualised as persistent depressive disorder (PDD). Unlike in the DSM-IV (APA, 2000), PDD has been introduced to include individuals who experience depressive symptoms over at least two years, but who do not meet MDD threshold (i.e., dysthymia). Chronic depression is associated with increased socio-economic disadvantages and higher comorbidity with other psychiatric conditions when compared to single-episode MDD (Murphy & Byrne, 2012).

Treatment-resistant Depression

Across the literature there is a lack of consensus on defining TRD. TRD is commonly defined as lack of response to two trials of antidepressant medication (ADM; Brown et al.,

2019), although in some trials TRD is used to describe patients with no response to one treatment attempt (e.g., Wiles et al., 2013). Alternative terms have been proposed to refer to TRD to decrease possible stigmatisation, such as 'refractory depression' or 'difficult to treat depression' (Demyttenaere & Van Duppen, 2019; McAllister-Williams et al., 2020). In terms of frequency, a larger-scale UK based study in primary care by Thomas et al. (2013) found that 55% of their sample presented with TRD (defined as lack of response to at least 6 weeks of adequate dosage of ADM). The personal and societal cost of TRD through cost of treatment, loss of work, increased suicidal risk, and caregiver burden is major when compared to non-treatment resistant depression (Demyttenaere & Van Duppen, 2019).

Recurrent Depression

Despite successful treatment for MDD, some studies suggest that 85% will experience a recurrence when followed-up over 15 years (Hardeveld et al., 2010; Mueller et al., 1999). As with chronic depression and TRD, recurrent depression poses major personal and societal costs (Greden, 2001).

In this review the terms TRD, chronic depression and recurrent depression will be used. Collectively all three subtypes of depression will be referred to as 'persisting forms of depression', emphasizing the challenges of achieving long-term, successful remission for a significant group of individuals with MDD. Such an approach has been taken in a review by McPherson and Senra (2022), who highlighted that there is much overlap between TRD, chronic depression and recurrent depression. Approximately 40% of patients with PDD are found to show treatment-resistance (Schramm et al. 2020). Additionally, lower treatment compliance can be an issue in the management of persisting forms of depression, contributing to relapse in chronic and recurrent depression (Gopinath et al., 2007). To support clinical practice in mental health services, it is important to understand these persisting forms of depression to improve treatment response and outcomes.

Psychological Treatment

The National Institute for Health and Care Excellence (NICE) guidelines (2022) recommend both pharmacological and psychological treatments for MDD. Psychological treatments for severe depression include cognitive behavioural therapy (CBT), counselling, interpersonal psychotherapy (IPT), and short-term psychodynamic psychotherapy (NICE, 2022). Furthermore, NICE guidance acknowledges that some patients may not respond to adequate treatment attempts. In those instances, a different ADM and/or psychological therapy should be tried.

To improve treatment outcomes, Cognitive Behavioural Analysis System of Psychotherapy (CBASP) has been developed specifically for chronic depression (McCullough, 2000), and was found more effective than non-specific psychological therapies (Schramm et al., 2017). Systematic reviews on psychological interventions for chronic depression and TRD provided further support for the benefit of psychological treatments (Cuijpers et al., 2010; Ijaz et al., 2018; Li et al., 2018; McPherson et al., 2005; Van Bronswijck et al., 2018). Interventions reviewed included CBASP, cognitive therapy (CT), CBT and IPT. To further improve outcomes, combination with ADM is indicated (Cuijpers et al., 2010; Ijaz et al., 2018). In relation to recurrent depression, research has focused on exploring which psychological interventions, such as CBASP, CBT, CT, IPT and mindfulness-based therapies can reduce recurrence (Biesheuvel-Leliefeld et al., 2015).

Despite evidence of the benefit of psychological interventions, over half of patients do not respond to the first pharmacological or psychological treatment they are offered (Van Bronswijck et al., 2018; Torpey & Klein, 2008). Given the wider impact of MDD, further research into psychological treatments of depression is encouraged (Cuijpers et al., 2017).

Although historically it has been argued that all psychological therapies are of similar efficacy (Luborsky et al., 1975), it is recognised that not every individual will benefit equally from all available psychotherapies. Evidence from meta-analyses of clinical trials indicate that, although the efficacy of different psychotherapies for the treatment of depression treatments is comparable (Palpacuer et al., 2017), patients may show better responses to one psychological treatment over another (Kaiser et al., 2022). This may explain why many individuals show no response to the first psychological treatment they are offered (Van Bronswijck et al., 2018). Given that NICE guidelines recommend a wide range of psychological approaches for MDD treatment, treatment allocation is often influenced by the involved professional's clinical judgement. Thereby, risk of bias during the decision-making process is often increased (Bell & Mellor, 2009; Hannan et al., 2005). Instead, increased understanding of baseline patient characteristics that predict or moderate treatment outcomes could provide clinicians with an objective approach to treatment allocation, and thus result in improved outcomes.

Predictors and Moderators

Predictors and moderators are variables measured at baseline assessments, which are associated with treatment outcomes. Kraemer et al. (2002) explain that "moderators specify for whom and under what conditions treatment works" (Kraemer et al., 2002, p. 878). This suggests that the interaction between a moderator variable and treatment type affects outcome. As a result, moderators allow clinicians to make informed decisions on which treatment an individual patient is most likely to respond to. Predictors are defined as "a baseline measure that has a main effect on outcome but no interactive effect" (Kraemer et al., 2002, p. 880). In clinical practice, this translates to predictors being variables that are generally associated with treatment outcome, regardless of treatment type. Hence, predictors are general prognostic indicators, whereas moderators are treatment specific.

Predictors and Moderators in the Treatment of Depression

Potential predictors and moderators of treatment outcome have been extensively researched within the field of depression, both for pharmacological and psychological treatment approaches (e.g., see Tanguay-Sela et al., 2022; Papakostas et al., 2022). However, reviews into predictors and moderators of treatment outcomes for more persisting forms of depression, such as TRD, chronic, and recurrent depression, are limited. For TRD specifically, predictors and moderators of treatment response have only been reviewed in relation to pharmacological treatment approaches (De Carlo et al., 2016). Some of the predictors of lower response rates to pharmacological treatment included: older age, higher numbers of past hospital admissions, presence of anxiety disorders, presence of personality disorders, current suicidal risk (De Carlo et al., 2016). Additionally, there is emerging evidence that certain factors, such as the experience of childhood trauma, can negatively affect psychological treatment response (Schramm et al., 2020). However, no existing reviews on variables that predict psychological treatment response for adults with persisting forms of MDD were found.

Aims

To the knowledge of the author, this is the first systematic review examining predictors and moderators of change following psychological interventions for persisting forms of depression. This review focuses on TRD, chronic depression and recurrent depression, which as previously discussed are all associated with poorer outcomes. The main objective is to explore predictors and moderators of change following psychological interventions fo persisting forms of depression. This is an important step in moving towards a personalised medicine approach, whereby treatment selection is guided by individuals baseline characteristics to increase chances of successful outcomes and reduce the high rate of treatment non-response in this clinical population (Kessler, 2018; Simon & Perlis, 2010).

Method

The Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement was used as a guideline for this review (Appendix A; Page et al., 2021).

Study Protocol

Scoping searches were conducted on the proposed review during October 2022. No published or ongoing reviews on the same topic were found. Subsequently a review protocol was registered with PROSPERO in November 2022 (registration number CRD42022379257).

Search Strategy

Three databases were systematically searched for relevant publications: Web of Science, PsycINFO and Scopus. Studies published between database inception and search date (1st of December 2022), which met pre-defined inclusion and exclusion criteria, were sought. Additionally, search alerts informing the author of new publications added to databases were set up for the 1st of December 2022 until the 13th of April 2023. All newly added studies were screened for suitability. Studies in both the English and German language were sought, due to the author being fluent in both languages. Titles, abstracts, and indexes were searched across the databases using the search strategy outlined in Table 1.

Table 1

Search Strategy

Construct	Search Term
Intervention	"psycho* therap*" OR "psycho* intervention" OR "cognitive behaviour* therap*" OR CBT or "cognitive therap*" OR counselling
Clinical Population	"treatment resistant depression" OR "chron* depress*" OR "recu* depress*" OR "relap* depress*" OR "persist* depress*"

Note. Constructs were combined using the Boolean operators OR / AND. The Boolean operator * was used to include varying endings of the given search term.

The search strategy was informed by reviews on similar patient and intervention groups (Ijaz et a., 2018; McPherson & Senra, 2022). Search terms for predictors and moderators were not included as this was found to narrow the search results and to increase the risk of missing studies which looked at predictors and moderators within secondary analyses. Draft search terms were trialled to ensure feasibility. The final search strategy was discussed with the research supervisor and a librarian. Although the registered review protocol stated that the grey literature would be searched, this was later decided against. This was due to time constraints, and the author noticing after database searches that studies of interest were primarily secondary data analyses to published randomised controlled trials (RCTs). Thus, focusing on published literature alone was considered to likely identify existing studies of interest. Additionally, published literature helps identify studies with larger sample sizes (Pappas & Williams, 2011), which are more likely to have completed predictor or moderator analyses.

Study Eligibility

The Population, Intervention, Comparison, Outcomes and Study (PICOS) framework was used to develop the research question and resulting inclusion and exclusion criteria (Table 2; Eriksen et al., 2018). PICO(s) frameworks are recognised to be beneficial search strategy tools for systematic reviews (Eriksen et al., 2018). Three depression subtypes were focused on, aimed at capturing the clinical population presenting with persisting depression symptomatology. TRD was defined as current MDD with at least one unsuccessful psychological or pharmacological intervention for depression. Due to the lack of consensus on defining TRD, an inclusive definition was chosen to allow review of a wider range of publications on TRD (Berlim & Turecki, 2007). Chronic depression was defined as the presence of depression symptoms for at least two years or recurrence/ relapse of depression symptoms during this time. This definition allows for inclusion of studies focusing on PDD as defined by the DSM-V (APA, 2013). Recurrent depression was defined as an individual having a diagnosis of MDD at the time of beginning a psychological intervention, in addition to at least one past episode of depression (APA, 2013).

Table 2

Inclusion and Exclusion Criteria

	Inclusion Criteria	Exclusion Criteria
Population	Patients (aged ≥ 18 years) with a diagnosis of major depressive disorder (in accordance with ICD-10 or DSM-V criteria or at the time of study valid diagnostic criteria). Diagnosis of major depressive disorder identified via diagnostic interview or by scoring above clinical threshold on a validated screening measure. Diagnosis meets the review criteria for treatment-resistant, chronic and/or recurrent depression.	Patients aged < 18 years. No formal diagnosis of major depressive disorder/ not scoring above clinical threshold on a validated screening measure. Diagnosis is not classed as treatment-resistant, chronic, or recurrent depression.
Intervention	Psychological intervention for depression, this can be combined with a pharmacological intervention.	Intervention is not based on a psychological model.
Comparator	Not applicable.	Not applicable.
Outcome	Standardised measure of depression symptoms (e.g., PHQ-9, HRSD, BDI-II), administered at least at baseline and post-intervention. Impact of at least one variable on post-intervention depression symptoms is statistically analysed.	Depression symptoms not measured using a standardised measure. Depression symptoms not measured at baseline and post-intervention. No statistical analysis of at least one variable on post-intervention depression symptoms.
Setting	Any outpatient setting.	Any inpatient setting.
Study Design	Randomised controlled trials and cohort studies. Written in English or German language.	Qualitative research, grey literature, conference proceedings, presentations, or popular media articles. Not written in English or German language.

Note. Beck Depression Inventory (BDI-II), Diagnostic and Statistical Manual of Mental Disorders 5th Edition (DSM-V), Hamilton Rating Scale for Depression (HRSD), International Classification of Diseases 10th revision (ICD-10), Patient Health Questionnaire-9 (PHQ-9).

Study Selection

Search results from all three databases were combined. Following removal of duplicates, articles were screened by title and abstract. Subsequently, full text articles were retrieved and screened against the inclusion criteria. Additionally, backward and forward citation searches, as well as searching of relevant trials and reviews, was completed. A second reviewer (KA) was given a random selection of studies identified for full text screen (n = 10). There was no disagreement in study selection between the author and second reviewer.

Data Extraction and Synthesis

Data extraction and narrative synthesis was informed by published guidance (Boland et al., 2014) and relevant reviews (Amati et al., 2017; Ijaz et a., 2018; McPherson & Senra, 2022). Extraction of relevant data (study characteristics, methodology, sample characteristics, psychological intervention, predictor/ moderator analyses and outcomes) was completed by the author. Predictors and moderators were defined in accordance with Kraemer et al. (2002). For secondary data analysis publications, the main trial publication was sought for data extraction where needed. Information was summarised in tables, with information from studies utilising the same original dataset clustered together. Due to the range of examined predictor and moderator variables, these were categorised into relevant groups and a narrative synthesis was conducted.

Risk of Bias Assessment

Risk of bias was assessed using the Critical Appraisal Skills Programme (CASP) case-control, cohort study or randomized controlled trial checklists (Appendices B to D; CASP, 2022). For secondary data analysis publications, the main trial publication was sought to complete the risk of bias assessment.

The author conducted risk of bias assessments on all studies. Between two independent second reviewers (LE, CG) all studies were once more assessed for risk of bias. Inter-rater reliability, calculated and interpreted using the Kappa statistic (Cohen, 1960; Landis & Koch, 1977), was extremely high between the author and both raters, $\kappa = .960$ ($SE = .020$, $p < .001$). Disagreement in ratings were discussed and a consensus was reached without the need of an additional reviewer.

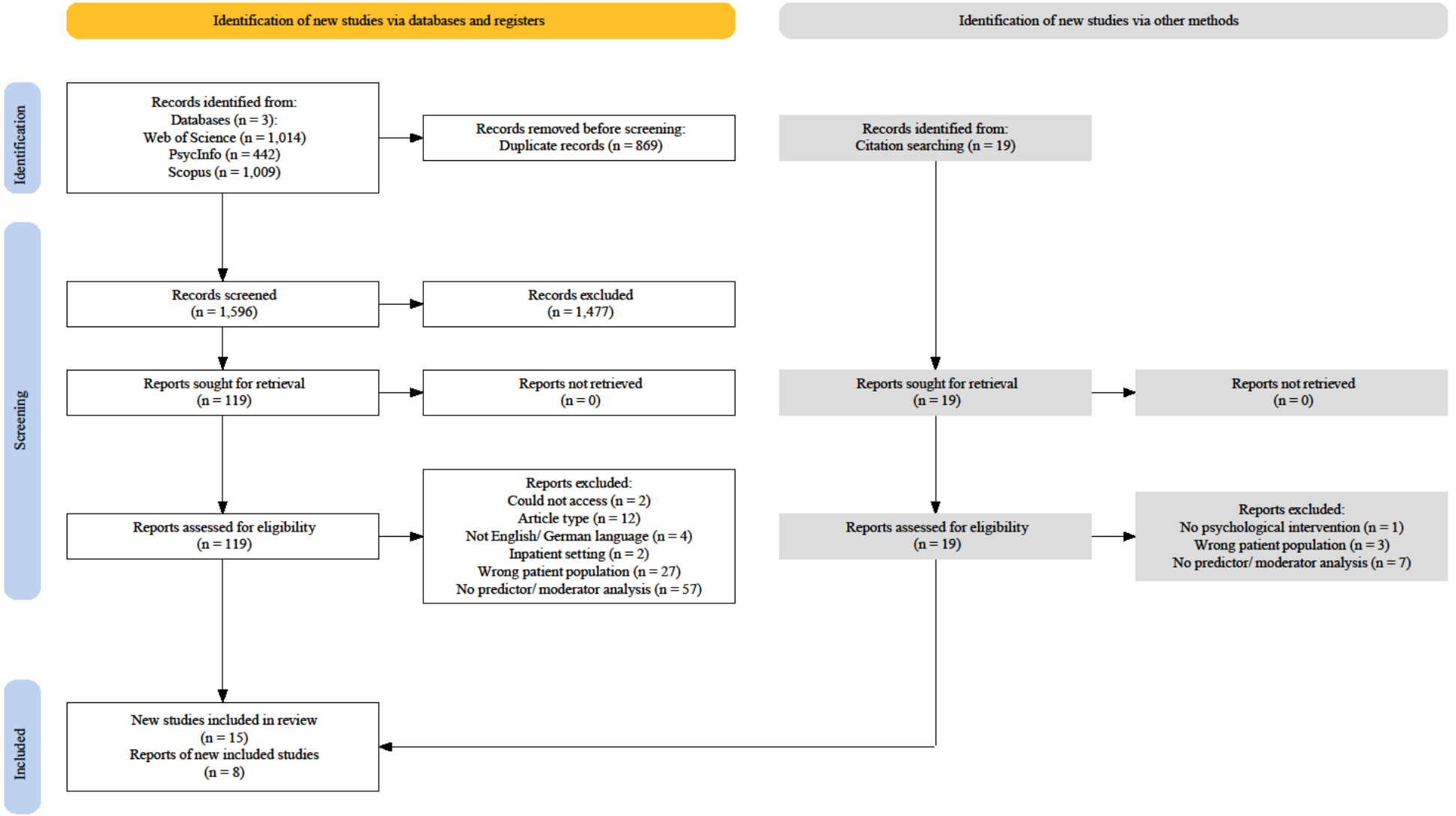
Results

Search Results

The study selection is outlined in Figure 1 using a PRISMA diagram (Page et al., 2021). Following the database searches, studies from all three databases were combined (n = 2465). Duplicates were removed (n = 869) and the remaining articles (n = 1596) were screened by title and abstract for relevance. For the remaining articles (n = 119), full text articles were retrieved and screened against the inclusion criteria. Reasons for excluding studies (n = 104) are summarised in Appendix E. Database searches yielded 15 studies for the systematic review. Citation searches led to identification of a further eight studies. A total of 23 studies were included in the review.

Figure 1

PRISMA Diagram Outlining Study Selection Process



Study and Participant Characteristics

An overview of study and participant characteristics is shown in Table 3. Out of the included studies, the majority were RCTs (*n* = 20). Most studies were conducted in the United States of America (*n* = 10), and the remaining in Europe. Sixteen studies were secondary data analyses to six other RCTs. Analysed sample sizes were large (≥ 100 participants) in 18 studies. The majority examined chronic depression (*n* = 16), with a further two studies looking at patients with co-occurring chronic depression and TRD. Four studies looked at TRD only, and one study at recurrent depression. All studies assessed presence of a depressive disorder utilising, at the time of participant recruitment, widely accepted diagnostic criteria, namely the ICD-10, DSM-IV, and DSM-V (APA, 2000; APA, 2013; World Health Organization, 1993). Eighteen studies utilised additional measures of depressive symptoms at assessment, with the majority (*n* = 15) using the clinician administered Hamilton Depression Rating Scale (HRSD).

The most frequent psychological intervention offered to participants was CBASP (*n* = 14). The remaining interventions offered were CBT (*n* = 3), mindfulness-based cognitive therapy (MBCT; *n* = 3), long term psychoanalytic psychotherapy (LTPP; *n* = 2), and MBCT combined with CBT (*n* = 1). Psychological interventions were primarily delivered on an individual basis (*n* = 18). Only one study did not have a control condition. The remaining studies had varying control conditions, namely ADM (*n* = 7), other psychological or psychosocial interventions (*n* = 11), combination treatments (*n* = 5), treatment-as-usual alone (*n* = 5), wait-list control (*n* = 2) and healthy controls (*n* =1). Ten studies had two control conditions, which included Renner and Berry (2011), Probst et al. (2020), five secondary data analyses to Keller et al. (2000), and two secondary data analyses to Kocsis et al. (2009b), . The third secondary data analysis to Kocsis et al. (2009b) excluded the ADM control condition (Arnow et al., 2013). Overall, the psychological interventions entailed eight to sixty sessions over an eight-to-sixty-week period. The use of ADM was allowed in all conditions or in the

ADM-only conditions for the majority of studies (*n* = 16). Overall, psychological interventions were found effective in achieving symptom reduction. Further detail on intervention characteristics is summarised in Appendix F. A variety of outcome measures were used to assess possible predictors and moderator variables, with the majority of studies (*n* = 15) utilising clinician-rated measures of depression.

Across all studies the majority of the sample were female, and where ethnicity was reported, predominantly of White ethnicity. In studies that provided information on past failed treatments, past failed interventions were reported for the majority of the sample indicating that a significant proportion of participants met the review's criteria of TRD.

Table 3

Overview of Study Characteristics

Study/ *Country*	**Study Design**	**Target Population/ TRD Definition**	**Assessment Criteria**	**Intervention**	**Comparison Condition(s)**	***N* (Total/ Analysed)**	**Age (years): *M*(SD)**	**Gender (% female)**	**Ethnicity (% White)**	**Past Failed Treatments**	**Main Depression Outcome**
Cladder-Micus et al. (2018) *Netherlands*	RCT	TRD (no response to ≥ 1 ADM + ≥1 psychological intervention)	DSM-IV; IDS-SR 21+	MBCT (group)	TAU	106/106	47.1 (10.25)	62%	NR	100%	IDS-SR
Eisendrath et al. (2016) *USA*	RCT	TRD (≥2 failed adequate ADM trials)	DSM-IV; $HDRS_{17}$ 14+	MBCT (group)	HEP	173/173	MBCT: 47.1 (13.46); HEP: 45.2 (11.19)	76%	80%	MBCT: *M*=2.9 HEP: *M*=3.06	QUIDS-SR
Lopez & Basco (2015) *USA*	Case Control Study	TRD (≥2 failed adequate ADM trials)	DSM-IV; QUIDS-SR 11+	CBT	TAU	166/166	43.1 (12.8)	84%	51%	100%	QIDS-SR
Potijk et al. (2020) *Netherlands*	Retro-spective Chart Review	Chronic AND TRD (≥2 failed treatments including ≥1 ADM)	DSM-IV	CBASP	None	54/54	Early onset cases: 50.8 (10.2); Late onset cases: 52.3 (8.2)	65%	NR	Psychotherapy: 98 % AD: 93 %; ECT: 11%; inpatient: 33%	IDS-SR
Renner & Berry (2011) *Austria*	RCT	Recurrent	ICD-10	CBT (group)	Self-Help Group / WL	66/34	42.7 (8.7)	100%	0%	NR	CES-D

Table 3 (continued)

Study/ *Country*	**Study Design**	**Target Population/ TRD Definition**	**Assessment Criteria**	**Intervention**	**Comparison Condition(s)**	***N* (Total/ Analysed)**	**Age (years): *M*(SD)**	**Gender (% female)**	**Ethnicity (% White)**	**Past Failed Treatments**	**Main Depression Outcome**
Stangier et al. (2021) *Germany*	RCT	Chronic	DSM-V	MBT (group) + CBT (individual)	WL	48/48	MBT (51.58 (11.26); WL (48.92 (11.39)	75%	NR	NR	QUIDS-C
Taubner et al. (2011) *Germany*	Case Control Study	Chronic	DSM-IV	LTPP	Healthy Controls	40/40	LTPP: 39.2 (12.7) Controls: 37.1 (11.6)	80%	NR	Psychotherapy and/or ADM: 80%	BDI-II
Secondary Data Analysis to Fonagy et al. (2015):											
Rost et al. (2019) *UK*	RCT	Chronic and TRD (≥2 failed adequate treatments including ≥1 ADM)	DSM-IV; $HRSD_{17}$ 14+; BDI-II 21+	LTPP	TAU	129/120	44.0 (10.31)	63%	81%	LTPP: *M*=3.5 (*SD*=1.4) TAU: *M*=3.9 (*SD*=1.8)	$HRSD_{17}$
Secondary Data Analysis to Keller et al. (2000):											
Arnow et al. (2003) *USA*	RCT	Chronic	DSM-IV; $HRSD_{24}$ 20+	CBASP	ADM / ADM + CBASP	681/347	44.9 (10)*	65% *	92% *	80% (Psychotherapy and/or ADM)	$HRSD_{24}$
Denton et al. (2010) *USA*						681/171	42.8 (9.0) *	64% *	91% *	80% (Psychotherapy and/or ADM)	IDS-SR_{30}

Table 3 (continued)

Study/ *Country*	**Study Design**	**Target Population/ TRD Definition**	**Assessment Criteria**	**Intervention**	**Comparison Condition(s)**	***N* (Total/ Analysed)**	**Age (years): *M*(SD)**	**Gender (% female)**	**Ethnicity (% White)**	**Past Failed Treatments**	**Main Depression Outcome**
Secondary Data Analysis to Keller et al. (2000) – continued:											
Kocsis et al. (2009a) *USA*						681/429	45 (NR)	65% *	92% *	ADM: 58% *	$HRSD_{24}$
Manber et al. (2008) *USA*	RCT	Chronic	DSM-IV; $HRSD_{24}$ 20+	CBASP	ADM / ADM + CBASP	681/681	43.5 (10.7)	65%	91%	80% (Psychotherapy and/or ADM)	$HRSD_{24}$
Nemeroff et al. (2003) *USA*						681/681	43.5 (10.7)	65%	91%	80% (Psychotherapy and/or ADM)	$HRSD_{24}$
Secondary Data Analysis to Phase 2 of Kocsis et al. (2009b):											
Arnow et al. (2013) *USA*					BSP	491/224	CBASP: 45.6(11.3)*; BSP: 47.4 (11.2)*	**CBASP:** 54%* **BSP:** 53%*	**CBASP**: 93% * **BSP**: 89% *	NR	$HRSD_{24}$
Schankman et al. (2013) *USA*	RCT	Chronic	DSM-IV; $HRSD_{24}$ 20+	CBASP	BSP / ADM	491/491	44.2 (1.2)	56%	88%	NR	$HRSD_{24}$
Steidtmann et al. (2012) *USA*					BSP / ADM	491/473	44.2 (1.2)	56%	88%	NR	$HRSD_{24}$

Table 3 (continued)

Study/ *Country*	**Study Design**	**Target Population/ TRD Definition**	**Assessment Criteria**	**Intervention**	**Comparison Condition(s)**	***N* (Total/ Analysed)**	**Age (years): *M*(SD)**	**Gender (% female)**	**Ethnicity (% White)**	**Past Failed Treatments**	**Main Depression Outcome**
Secondary Data Analysis to Michalak et al. (2015):											
Probst et al. (2020) *Germany*	RCT	Chronic	DSM-IV	MBCT (group)	CBASP (group)/ TAU	106/68	MBCT: 48.09 (11.62)*; CBASP: 51.03, (10.60)*	62% *	NR	NR	$HRSD_{24}$
Secondary Data Analysis to Schramm et al. (2017) and Schramm et al. (2019):											
Assmann et al. (2018) *Germany*	RCT	Chronic	DSM-IV; $HRSD_{24}$ 20+	CBASP	SP	268/268	44.9 (11.8)	66%	NR	psychotherapy: 57%; ADM: 55%; Combination: 20%	$HRSD_{24}$
Bausch et al. (2020) *Germany*						268/256	44.9 (11.8)	66%	NR	psychotherapy: 57%; ADM: 55%; Combination: 20%	$HRSD_{24}$ + IDS-SR
Erkens et al. (2018) *Germany*						268/247	44.9 (11.8)	66%	NR	psychotherapy: 57%; ADM: 55%; Combination: 20%	$HRSD_{24}$
Klein et al. (2018) *Germany*						268/256	44.9 (11.8)	66%	NR	psychotherapy: 57%; ADM: 55%; Combination: 20%	$HRSD_{24}$

Table 3 (continued)

Study/ *Country*	Study Design	Target Population/ TRD Definition	Assessment Criteria	Intervention	Comparison Condition(s)	*N (*Total/ Analysed)	Age (years): *M*(SD)	Gender (% female)	Ethnicity (% White)	Past Failed Treatments	Main Depression Outcome
Secondary Data Analysis to Schramm et al. (2017) and Schramm et al. (2019) – continued:											
Serbanescu et al. (2020) *Germany*	RCT	Chronic	DSM-IV; $HRSD_{24}$ 20+	CBASP	SP	268/268	44.9 (11.8)	66%	NR	psychotherapy: 57%; ADM: 55%; Combination: 20%	$HRSD_{24}$
Secondary Data Analysis to Wiles et al. (2013):											
Button et al. (2015) *UK*	RCT	TRD (no response to ≥1 ADM taken for at least 6 weeks)	ICD-10; BDI-II 14+	CBT	TAU	469/469	CBT: 49.2 (11.9); TAU: 50 (11.5)	72%	99%	ADM: 80%	BDI-II

Note. Participant demographic and clinical characteristics of participants provided for total sample where available. * Denotes that information refers to the analysed sample only.

Antidepressant medication (ADM), Becks Depression Inventory II (BDI-II), Brief Supportive Psychotherapy (BSP), Cognitive Behavioural Analysis System of Psychotherapy (CBASP), Diagnostic and Statistical Manual of Mental Disorders 4th Edition or 5th Edition (DSM-IV, DSM-V), Hamilton Depression Rating Scale 17 item or 24 item version ($HDRS_{17}$, $HDRS_{24}$), Health Enhancement Programme (HEP), International Classification of Diseases 10th revision (ICD-10), The Inventory of Depressive Symptomatology Self-Report (ISD-SR), Long Term Psychoanalytic Psychotherapy (LTPP), Mindfulness-based Cognitive Therapy (MBCT), Metta-Based Therapy (MBT), Not reported (NT), Quick Inventory of Depressive Symptomatology Self-Report or Clinician-Rated (QUIDS-SR, QUIDS-C), Randomised Controlled Trial (RCT), Supportive Psychotherapy (SP), Treatment as usual (TAU), Treatment Resistant Depression (TRD), Wait List (WL).

Risk of Bias Assessment

Out of all included RCTs, nine were considered to have a low, ten a medium and one study a high risk of bias. Across all studies participants and therapists were not blinded to the allocation condition. No study completed a power analysis specifically for the predictor or moderator analysis. Other common reasons for increased risk of bias included: small sample sizes restricting generalisability of findings, insufficient information on results such as missing p-values or partial reporting of results for all measures for each timepoint, and some differences in baseline characteristics of participants across the conditions.

For the two case control studies, one was rated as low risk of bias (Lopez & Basco, 2015) and one as high risk of bias (Taubner et al., 2011). Taubner et al. (2011) utilised a healthy control group, affecting validity of findings. Potijk et al. (2020), a cohort study, was rated as having a medium risk of bias. This is due to lack of follow-up and lack of clarity on treatment fidelity. All reviewed studies utilised valid and reliable measures to assesses outcomes and examined predictor or moderator variables. Further detail on all risk of bias assessments is summarised in Appendix G.

Narrative Synthesis – Predictors and Moderators

Across all reviewed studies, 65 different variables were analysed as potential predictors or moderators. An overview of examined variables, statistical analyses, and key findings is outlined in Table 4. To aid interpretation of findings, variables were categorised as follows: sociodemographic characteristics, clinical characteristics, interpersonal and personality factors, psychological factors, and treatment factors. Clinical characteristics were divided into four further sub-categories: depression characteristics, baseline clinical characteristics, comorbidities, and trauma factors.

Two studies utilised cross-validation approaches in their analysis of potential predictors and moderators (Manber et al., 2008; Serbanescu et al., 2020). Manber et al. (2008) used Receiver operating characteristics curve (ROC) analysis, where once a significant predictor and cut-off point is identified the sample is divided into two subgroups and predictor testing is re-started for each subgroup separately. Serbanescu et al. (2020) used cross-validation methods to calculate a composite moderator M* which included all tested potential moderator variables with an effect size ≥ 0.10.

Table 4

Summary of Key Findings of Predictor and Moderator Analyses

Study	Statistical Analysis	Variable (Measure)	Key Findings
Cladder-Micus et al. (2018)	ANCOVA	Age	NS
		Baseline depression score (IDS-SR)	NS
		Childhood trauma (CTQ)	NS
		Duration current episode	NS
		Gender	NS
		Mindfulness (FFMQ)	NS
		N previous episodes	NS
		Rumination (RRS)	Moderator, F (1, 84) = 5.44, p =0.02
		Self-compassion (SCS)	NS
		Treatment resistance (DM-TRD)	NS
Eisendrath et al. (2016)	Multivariate Regression	Age at onset	NS
		Baseline anxiety (STAI)	Predictor **
		Childhood trauma (CTQ)	Predictor **
		Current Stress (PSS)	Predictor **
		Disability status	NS
		Duration current episode	NS
		Education	NS
		Ethnicity	NS
		Medical illness	NS
		Minority and socioeconomic status	NS
		N previous episodes	NS
		Personality disorder (SCID)	Predictor **
Lopez & Basco (2015)	Multilevel Regression	Age	Predictor, b = .00, SE = .00, t = 2.46, p = .014
		Gender	Predictor, b = .05, SE = .02, t = -2.14, p = .033
		Ethnicity	NS
		Marital status	NS
		Baseline depression score (QIDS-SR)	Predictor, b = -.00, SE = .00, t = -2.11, p = .036
		Past inpatient treatment	Predictor, b = -.03, SE = .02, t = -2.11, p = .046
		Personality disorder	NS
		Substance-related disorder	NS
Potijk et al. (2020)	Independent t-test, Chi-Square test	Age at onset	Predictor, p = .010
Renner & Berry (2011)	Linear Regression	Age	Predictor, b = -.05, SE = .02, p = .004
		Duration of stay in Austria	Predictor, b = .03, SE = .01, p = .034
		Education	NS
		Generation of migration	NS
		Number of children	NS
		Traumatic events experienced (LEC)	Predictor, b = .09, SE = .03, p = .004
		Traumatic events witnessed (LEC)	NS

Table 4 (continued)

Study	Statistical Analysis	Variable (Measure)	Key Findings
Stangier et al. (2021)	MANOVA	Compassion to others (CLS)	NS
Taubner et al. (2011)	Multiple Hierarchical Regressions	Reflective functioning (RFS)	NS
Secondary Data Analysis to Fonagy et al. (2015):			
Rost et al. (2019)	Multilevel Regression Model	Personality features (AIDA)	Moderator, b = -.91, SE = 0.44, p = .038
Secondary Data Analysis to Keller et al. (2000):			
Arnow et al. (2003)	Multiple Regression Model	Baseline depression scores ($HRSD_{24}$) Gender Therapeutic reactance (TRS) *'Inner Directed' Factor* *'Defiant/Oppositional' Factor*	Predictor, B = .20, p = .037 NS Predictor b = .22, p = .04 b = .27, p = .012
Denton et al. (2010)	Logistic regression	Dyadic discord (MAS)	Predictor, $\chi 2$ = 8.8, df = 1, p = 0.003
Kocsis et al. (2009a)	Logistic regression	Treatment preference	Moderator, $\chi 2$ = 13.29, df = 6, p = .039
Manber et al. (2008)	ROC analysis	Age Gender Ethnicity Marital Status Employment Status Baseline depression score ($HRSD_{24}$) Baseline anxiety score (HAM-A) Age at onset Duration current episode Childhood trauma (CTS) Attributional style (ASQ) Social functioning (SAS) Treatment group	NS NS NS NS NS Predictor, p < .01 Predictor, p < .01 NS NS NS NS NS Predictor, $\chi 2$ = 19.7, p < .001
Nemeroff et al. (2003)	Linear regression & LOCF analysis	Childhood Trauma (CTS)	Moderator, OR = 2.322, 95% CI =1.225–4.066
Secondary Data Analysis to Phase 2 of Kocsis et al. (2009b):			
Arnow et al. (2013)	Linear Mixed Regression	Age Gender Global functioning (GAF)	Predictor, F(1, 609) = 3.72, p = .05 NS Predictor, F(1, 1218) = 84.85, p < .001

Table 4 (continued)

Study	Statistical Analysis	Variable (Measure)	Key Findings
Secondary Data Analysis to Phase 2 of Kocsis et al. (2009b) – continued:			
Schankman et al. (2013)	Linear Mixed Regression	Dysfunctional Attitudes (DAS)	Moderator, b = -.0009, t (285) = - 3.19
Steidtmann et al. (2012)	Linear Mixed Regression	Treatment Preference	NS
Secondary Data Analysis to Michalak et al. (2015):			
Probst et al. (2020)	Multilevel regression model	Interpersonal Problems (IIP-32) *'vindictive/self-centred' subscale* *'non-assertive' subscale*	Moderator Estimate = 5.12 (*SE* = 1.71); $p < .01$ Estimate = -9.14 (*SE* = 2.84); $p < .01$
Secondary Data Analysis to Schramm et al. (2017) and Schramm et al. (2019):			
Assmann et al. (2018)	ANCOVA	Anxiety Disorder (SCID)	Moderator, $F_{1,256} = 7.06$, $p = 0.01$
Bausch et al. (2020)	Linear Mixed Model	Childhood Trauma (CTQ)	NS
Erkens et al. (2018)	ANCOVA	Personality Disorder (SCID)	NS
Klein et al. (2018)	ANCOVA	Childhood Trauma (CTQ) *Emotional neglect subscale*	Moderator, $F(1, 244) = 4.253$, $p = 0.040$ Moderator, $F(1, 244) = 6.866$, $p = 0.009$ and Predictor, $F(1, 244) = 8.565$, $p = 0.004$.
Serbanescu et al. (2020)	Multi-variable and lasso regression with k-fold cross validation	Gender Age Being Single Married/ cohabiting Separated/divorced Widowed Education Being employed Presence of ≥ 1 morbidity	NS NS NS NS M*, $r = 0.34$ (95% CI, 0.32; 0.36) M*, $r = 0.34$ (95% CI, 0.32; 0.36) NS NS NS

Table 4 (continued)

Study	Statistical Analysis	Variable (Measure)	Key Findings
Secondary Data Analysis to Schramm et al. (2017) and Schramm et al. (2019) – continued:			
Serbanescu et al. (2020)	Multi-variable and lasso regression with k-fold cross validation	Chronic depression	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Double depression	NS
		Recurrent depression	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Age at onset	NS
		Duration current episode	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Baseline depression score (IDS-SR)	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Baseline depression score ($HRSD_{24}$);	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Suicidality (BSSI)	NS
		History of ≥ 1 suicide attempt	NS
		Baseline anxiety (GAD-7)	NS
		Baseline anxiety (BSI)	NS
		Baseline phobic anxiety (BSI)	NS
		Axis I disorder (SCID)	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Axis II disorder (SCID)	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Global functioning (GAF)	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Social functioning (SASS)	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Quality of life (QLDS)	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Interpersonal problems (IIP-32)	NS
		Childhood emotional abuse (CTQ)	NS
		Childhood emotional neglect (CTQ)	M*, r = 0.34 (95% CI, 0.32; 0.36)
		Childhood physical abuse (CTQ)	NS
		Childhood physical neglect (CTQ)	NS
		Childhood sexual abuse (CTQ)	NS
		Past treatment type	NS
		Past inpatient treatment	NS
		Treatment preference	NS
Secondary Data Analysis to Wiles et al. (2013):			
Button et al. (2015)	Random Effects Regression Model	Age	Moderator, b=0.24 (95% CI 0.44, 0.04), p = 0.02
		Baseline anxiety (CIS-R)	NS
		Baseline depression score (CIS-R)	NS
		Baseline depression score (BDI-II)	Moderator, b = 0.20 (95% CI 0.00, 0.39), p = 0.05
		Baseline PTSD score (PC-PTSD)	NS
		Current Stress (SRRS)	NS
		Duration current episode	NS
		Dysfunctional attitudes (DAS)	NS
		Education	NS
		Longstanding illness	NS
		Marital status	NS
		Meta-cognitive awareness (MAQ)	NS
		N previous episodes	NS
		Neuroticism (BFI)	NS

Note. M* - Composite Moderator Score calculated using variables with an effect size of d ≥ 0.10 regardless of significant level. ** significance levels were unavailable.

Anaclitic-Introjective-Depression Assessment (AIDA), Analysis of covariance (ANCOVA), Attributional Style Questionnaire for Negative Events (ASQ), Beck Depression Inventory (BDI-II), 'Big Five' Inventory (BFI)', Brief Symptom Inventory (BSI), Beck Scale for Suicide Ideation (BSSI), Clinical Interview Schedule -Revised (CIS-R), Compassionate Love Scale (CLS), Childhood Trauma Questionnaire (CTQ), Childhood Trauma Scale (CTS), Dysfunctional Attitude Scale (DAS), Dutch Measure for Quantification of Treatment Resistance in Depression (DM-TRD), Five Facets Mindfulness Questionnaire (FFMQ), General Anxiety Disorder-7 (GAD-7), Global Assessment of Functioning Scale (GAF), Hamilton Anxiety Scale (HAM-A), Hamilton Rating Scale for Depression (HRSD), Inventory of Depressive Symptomatology Self-Report (IDS-SR), Inventory of Interpersonal Problems (IIP-32), Life Events Checklist from the Clinical Administered PTSD Scale (LEC), Last observation carried forward (LOCF), Multivariate analysis of covariance (MANOVA), Meta-cognitive Awareness Questionnaire (MAQ), Marital Adjustment Scale (MAS), Non-Significant (NS), Primary Care PTSD Screening Tool (PC-PTSD), Perceived Stress Scale (PSS), Quick Inventory of Depressive Symptomatology Self-report (QIDS-SR), Quality of Life In Depression Scale (QLDS), Reflective Functioning Scale (RFS), Receiver operating characteristic analysis (ROC), Ruminative Response Scale (RRS), Social Adjustment Scale (SAS), Social Adaptation Self-evaluation Scale (SASS), Structured Clinical Interview (SCID-II), Self-Compassion Scale (SCS), Social Readjustment Rating Scale (SRRS), State-Trait Anxiety Inventory (STAI), Therapeutic Reactance Scale (TRS).

Sociodemographic Characteristics

Across studies 12 socio-demographic variables were analysed. Age was examined in seven studies with mixed findings. Three studies did not find age to be a predictor or moderator of treatment outcome (Cladder-Micus et al., 2018; Manber et al., 2008; Serbanescu et al., 2020). Button et al. (2015) found age to be a significant moderator, noting that higher age was associated with better treatment outcomes in CBT as opposed to treatment-as-usual (TAU). Three studies found age to be a significant predictor. Arnow et al. (2013) found younger age to predict lower post-intervention depression scores only for patients in the Brief Supportive Psychotherapy (BSP) control condition, but not in the CBASP condition. Lopez and Basco (2015) found younger participants to show faster response rates to CBT and TAU compared to older participants. Similar findings were reported by Renner and Berry (2011), who compared CBT to a structured self-help group (SHG).

Five studies did not find gender to be a significant predictor or moderator (Arnow et al., 2003; Arnow et al., 2013; Cladder-Micus et al., 2018; Manber et al., 2008; Serbanescu et al., 2020). In contrast, Lopez and Basco (2015) found gender to be a predictor of improvement rate, with female participants improving at a faster rate and showing greater benefit from CBT than male participants.

Marital status was not a significant predictor or moderator (Button et al., 2015; Lopez & Basco, 2015; Manber et al., 2008; Serbanescu et al., 2020). However, Serbanescu et al. (2020) additionally analysed 'being divorced/ widowed' and 'being separated' as separate variables, both of which met threshold for inclusion in the overall calculation of a composite moderator score. Level of education (Button et al., 2015; Eisendrath et al., 2016; Renner & Berry, 2011; Serbanescu et al., 2020), ethnicity (Eisendrath et al., 2016; Lopez & Basco, 2015; Manber et al., 2008), employment status (Manber et al., 2008; Serbanescu et al., 2020) and

minority and socio-economic status (Eisendrath et al., 2016) were not found to be predictors or moderators.

Renner and Berry (2011) studied treatment approaches for Turkish women with recurrent depression who migrated to Austria. Therefore, additional sociodemographic variables were tested as possible predictors: generation of migration, number of children, duration of stay Austria. Only duration of stay was a significant predictor, with greater number of years lived in Austria associated with better outcomes.

Clinical Characteristics

Baseline Clinical Characteristics. Baseline depression scores were consistently found to be predictors or moderators across all studies which examined this variable ($n = 5$) . Baseline depression was assessed using self-report (Beck Depression Inventory; Inventory of Depressive Symptomatology self-report; Quick Inventory of Depressive Symptomatology self-report) or clinician-rated outcome measures (HRSD, Clinical Interview Schedule - Revised). Three studies found that lower baseline depression levels were associated with better post-intervention outcomes (Arnow et al., 2003; Button et al., 2015; Manber et al., 2008; Lopez & Basco, 2015). In Manber et al. (2020) this was specific to those receiving CBASP. In Button et al. (2015) this was specific to the CBT sample and only when baseline depression was assessed using a self-report measure, but not using when using the clinician-rated measure. Serbanescu et al. (2020) found participants with higher baseline depression severity to benefit significantly more from CBASP than supportive psychotherapy (SP).

Baseline anxiety, as well as general and phobic anxiety, was not a significant predictor or moderator (Button et al., 2015; Manber et al., 2008; Serbanescu et al., 2020). However, Eisendrath et al. (2016) measured state and trait anxiety using the State-Trait Anxiety Inventory and found that state anxiety predicted smaller reductions in depression symptoms. Quality of

life had a moderating effect, with higher baseline quality of life showing better outcomes with SP than CBASP (Serbanescu et al., 2020). Likewise, participants with higher baseline general and social functioning responded better to SP than CBASP (Serbanescu et al., 2020). Arnow et al. (2013) found baseline global functioning to be a significant predictor, with higher baseline scores associated with lower post-intervention scores. Contrary to Serbanescu et al. (2020), Manber et al. (2008) did not find baseline social functioning to be a significant predictor or moderator. Eisendrath et al. (2016) examined impact of baseline stress levels, noting that higher scores predicted poorer outcomes. However, Button et al. (2015) assessed stress levels following adverse life events using the Social Readjustment Rating Scale and found a non-significant effect on outcomes. Baseline post-traumatic stress disorder levels (Button et al., 2015), baseline levels of suicidality and history of suicide attempts (Serbanescu et al., 2020) were not significant predictors or moderators.

Depression Characteristics. Seven different variables were examined. Serbanescu et al. (2020) explored depression type (chronic, double depression and recurrent depression) as potential moderators. Only double depression did not meet threshold for inclusion in the composite moderator variable. Chronic depression was associated with better response to CBASP, recurrent depression with better treatment response to SP. Number of previous episodes (Button et al., 2015; Eisendrath et al., 2016) and level of treatment resistance (Cladder-Micus et al., 2018) were not significant predictors or moderators. Age of depression onset was not found to be a significant predictor or moderator in three studies (Eisendrath et al., 2016; Manber et al., 2008; Serbanescu et al., 2020). In contrast, Potijk et al. (2020) found that patients with late-onset chronic depression (i.e., onset after 21 years of age) had significantly higher remission rates that those with early-onset chronic depression. However, this difference was not found when comparing pre- to post-intervention score changes on the Inventory of Depressive Symptomatology self-report measure (Potijk et al., 2020). Duration of

episode did not predict or moderate outcomes in two studies (Eisendrath et al., 2016; Manber et al., 2008), but did meet threshold for the composite moderator variable in Serbanescu et al. (2020) where longer episode duration was associated with better outcomes to CBASP as opposed to SP.

Comorbidities. Some comorbidities were found to be significant predictors or moderators. Serbanescu et al. (2020) found that participants with an Axis I comorbidity benefitted more from CBASP, whereas those with an Axis II disorder benefitted more from SP. For the same sample, an analysis by Assmann et al. (2018) showed that those with an anxiety disorder responded significantly better to CBASP than SP. Personality disorder presence was not associated with treatment outcome in two studies (Lopez & Basco, 2015; Erkens et al., 2018), but Eisendrath et al. (2016) found that presence predicted significantly worse outcomes. Presence of at least one morbidity (Serbanescu et al., 2020; Button et al., 2015), a substance-related disorder (Lopez & Basco, 2015), a disability (Eisendrath et al., 2016) or medical illness (Eisendrath et al., 2016) did not impact on outcomes.

Trauma Factors. There was some evidence to suggest that the experience of trauma can affect outcomes. Renner and Berry (2011) examined impact of lifetime traumatic events, witnessed and experienced, with only latter being a significant predictor. The higher the number of traumatic events experienced, the greater the benefit from CBT or the SHG was. Childhood trauma was assessed using the Childhood Trauma Questionnaire (CTQ) or Childhood Trauma Scale (CTS) across seven studies. Three did not find a significant relationship with treatment outcome (Bausch et al., 2020; Cladder-Micus et al., 2018; Manber et al., 2008). It is important to note that Bausch et al. (2020) compared baseline only with one- and two-year follow up depression scores. Klein et al. (2018) analysed data from the same trial as Bausch et al. (2020), however only focused on pre- and immediate post-intervention scores.

They found that overall presence of childhood trauma and childhood emotional neglect were moderators of treatment outcome, noting that CBASP should be preferred to SP. Additionally childhood emotional neglect was also a significant predictor, presence of which was associated with worse outcomes. Significant findings were reported by a further three studies. Eisendrath et al. (2016) found that only the experience of emotional abuse or neglect was predictive of poorer outcomes. Serbanescu et al. (2020) found only the experience of emotional neglect to be a moderator, noting that CBASP would be the preferred treatment compared to SP. Nemeroff et al. (2003) was the only study that used the CTS instead of the CTQ. Nemeroff et al. (2003) found childhood trauma, as well as the CTS subcategories of parental loss, physical abuse, and neglect, to be moderators of outcome. If childhood trauma was present, CBASP showed better outcomes than pharmacological treatment alone.

Interpersonal and Personality Factors

Overall, few studies assessed interpersonal and personality variables. Presence of relationship challenges, described as dyadic discord, was found to predict lower remission rates (Denton et al., 2010). Furthermore, interpersonal problems measured using the Inventory of Interpersonal Problems were moderators in Probst et al. (2020). Those scoring high on the 'vindictive/self-centred' subscale benefitted more from CBASP, whereas those scoring high on the 'non-assertive' subscale benefitted more from MBCT. However, moderating effects of interpersonal problems was not supported by Serbanescu et al. (2020). Rost et al. (2019) found that certain personality features assessed using the Anaclitic-Introjective-Depression assessment moderated treatment outcomes. Those with ‘self-critical’ or ‘needy’ features benefitted more from LTPP than TAU.

Psychological Factors

The majority of psychological factors examined were not found to predict or moderate outcomes. This included attributional style (Manber et al., 2008), compassion to others

(Stangier et al., 2021), dysfunctional attitudes (Button et al., 2015), meta-cognitive awareness (Button et al., 2015), mindfulness skills (Cladder-Micus et al., 2018), neuroticism (Button et al., 2015), reflective functioning (Rost et al., 2019) and self-compassion (Cladder-Micus et al, 2018). Only rumination was a moderator, with higher baseline rumination associated with larger decrease in depression symptoms in MBCT (Cladder-Micus et al., 2018).

Treatment Factors

Results provide limited evidence that treatment factors affect outcome. Therapeutic reactance was examined by Arnow et al. (2003) and was found to be a predictor of outcome for CBASP only. Those who had higher 'inner directed' or 'defiant oppositional' scores had higher depression symptom reduction. Treatment preference was not found to affect outcomes in two studies (Serbanescu et al., 2020; Steidtman et al., 2012), but did in Kocsis et al. (2009a). Kocsis et al. (2009a) reported that those receiving their preferred treatment had higher rates of remission and partial response. In terms of past treatment types, Serbanescu et al. (2020) did not find these to be moderators. However, Lopez & Basco (2015) found that past inpatient treatment was predictive of quicker symptom improvement.

Discussion

This systematic review examined potential predictors and moderators of response to psychological treatment for persisting forms of depression. A total of 23 studies examining 65 variables across five domains (sociodemographic, clinical, interpersonal/ personality, psychological and treatment variables) were included. Over half (57%) of examined variables were not found to be significant predictors or moderators. In some cases (25%), findings were inconclusive amongst reviewed studies. Eighteen percent of variables were found to be significant predictors or moderators but were mostly examined in individual studies, thus replication of findings was lacking. However, some variables were studied more often than others (in at least five studies), namely age, gender, baseline depression severity, and the

experience of childhood trauma. These sociodemographic and clinical factors are therefore more closely examined in the discussion. Findings are compared to the wider literature on MDD, which refers to research into all subtypes of MDD (including single-episode MDD) across the lifespan.

Contribution to the Evidence Base

Sociodemographic Characteristics

Gender was not found to be a predictor or moderator in most studies that examined the variable, all of which were RCTs with low or medium risk of bias. The only study that found gender to be a predictor was a case-control study (Lopez and Basco et al., 2015). However, due to case-control studies being more susceptible to the impact of confounding variables than RCTs (Tenny et al., 2017), this significant finding is viewed with caution. The conclusion that an individual's gender is an unlikely predictor or moderator is consistent with the wider literature. A review by Nilsen et al. (2013) into depression in children and adolescents did not find gender to be a predictor or moderator. Likewise, an individual patient data meta-analysis comparing CBT and ADM did not find gender to be a moderator or predictor (Cuijpers et al., 2014).

In contrast, findings for age as a potential predictor or moderator were mixed. The quality of studies who reported non-significant findings ranged from low to medium, and were all RCTs (Cladder-Micus et al., 2018; Manber et al., 2008; Serbanescu et al., 2020). Studies that found younger age to be associated with more favourable treatment outcomes were varied in quality, with one case-control and one high risk of bias study included (Lopez & Basco, 2015; Renner & Berry, 2011). When compared with wider MDD research, findings on age as a predictor remain inconclusive. Although some studies found younger age to be associated with better outcomes (Fournier et al., 2009), this is contrasted by reviews which report mixed findings across reviewed studies (Cuijpers et al., 2020; Nilsen et al., 2013). Cuijpers et al.

(2020) noted that age may be a relevant predictor for children and adolescents but found limited evidence to support age as a predictor across the adult lifespan. When considering the quality of studies examining age in this review, as well as the wider literature, caution should be taken when viewing age as a potential predictor or moderator.

Baseline Depression Severity

Baseline depression severity was consistently found to a be a predictor of treatment response, with lower baseline severity associated with better outcomes (Arnow et al., 2003; Button et al., 2015; Manber et al., 2008; Lopez & Basco, 2015). However, it should be noted that Arnow et al. (2003) and Manber et al. (2008) conducted secondary data analyses on the same trial data (Keller et al., 2000), therefore duplicating findings. Nevertheless, baseline depression severity is noted to be a robust predictor of outcomes across over the literature. Three reviews into depression note that lower baseline severity is predictive of better outcomes (Lorenzo-Luaces et al., 2017; Nilsen et al., 2013; Tunvirachaisakul et al., 2018). Driessen at al. (2010) also found baseline depression severity to predict outcome, but in the opposite direction. Adult outpatients with higher baseline severity appeared to benefit more than those with low baseline severity from psychological interventions. However, Driessen et al. (2010) noted that a relatively small number of cases with severe depression were included in their meta-analysis, whereas this review focuses on persisting forms of depression where depression severity is likely higher. Secondly, Driessen et al. (2010) focused on post-treatment effect sizes. It would be of interest to examine whether remission rates varied between low- and high-severity cases.

In terms of moderating effects, only Serbanescu et al. (2020) conducted a moderator analysis in this review. Higher baseline scores in chronic depression indicated CBASP as the favoured treatment over BSP. In contrast, a meta-analysis by Weitz et al. (2015) compared CBT versus ADM treatments for depression and did not find baseline depression severity to be

a moderator. However, it is important to note that Serbanescu et al. (2020) compared two psychological treatment approaches, CBASP and SP. One of these interventions, CBASP, was specifically developed for individuals with chronic depression and is found effective (Schramm et al., 2017; Cuijpers et al., 2010; Ijaz et al., 2018). It is possible that more severe cases of chronic depression benefitted more from a targeted intervention (CBASP) than from a non-directive approach (SP).

Given the consistency in findings that baseline depression severity is associated with treatment outcomes, there is some preliminary evidence to support baseline depression severity as a potential predictor or moderator for persisting forms of depression. Confidence is further increased given the low to medium risk of bias across studies. Similarly, the wider literature on depression repeatedly notes baseline severity to be associated with outcomes.

Childhood Trauma

Presence of childhood trauma was supported in some studies as a potential predictor or moderator, whereas not in others. All studies were RCTs with low to medium risk of bias, indicating acceptable study quality. However, several studies utilised the same original data sources, whilst showing different outcomes. A possible explanation for differences in findings for studies using the same participant data includes differences in chosen statistical analysis, namely regression analysis versus ROC analysis (Manber et al, 2008; Nemeroff et al., 2003). Secondly, one study focused on long-term post-intervention outcomes (Bausch et al., 2020) as opposed to immediate post-intervention outcomes (Klein et al., 2018; Serbanescu et al., 2020).

Consistent with findings of this review, wider research into the impact of childhood trauma on treatment outcomes is mixed. Given that childhood trauma is a well-recognised risk factor for the development of MDD, a recently published meta-analysis by the Childhood Trauma Meta-Analysis Study Group (2022) explored whether it was also associated with

differential treatment outcomes. No significant differences in outcome were found between those with and without childhood trauma. This is in contrast to Nanni et al. (2012) and Nelson et al. (2017) whose meta-analyses found that presence of childhood trauma was predictive of poor outcomes. In conclusion, it is unclear whether childhood trauma is a robust predictor or moderator of outcomes in persisting forms of depression. Further research into the potential predictive and moderating effects of childhood trauma should be considered.

Other Baseline Characteristics

Several baseline characteristics were not significant predictors or moderators. However, non-significant predictors and moderators might challenge clinicians on their own preconceptions on who is and is not likely to benefit from an intervention. For instance, several socio-demographic characteristics were repeatedly not found to be predictors or moderators, such as education level, ethnicity, and marital status. This is consistent with wider MDD research, where demographic characteristics are generally not found to be useful criteria for treatment allocation (Sharpley & Bitsika, 2011). Baseline anxiety was examined by several studies in this review but was mostly not found to be a predictor or moderator of treatment outcomes. This is in contrast to the wider MDD literature, where baseline anxiety has been found to be a predictor of outcome (Kiosses et al., 2011; Papakostas et al., 2022; Tunvirachaisakul et al., 2018). Additionally, the presence of co-morbidities is often considered to be predictive of worse outcomes in MDD (Tanguay-Sela et al., 2022; Tunvirachaisakul et al., 2018). However, this review showed there is insufficient evidence to support this to be the case for persisting forms of depression.

Methodological Considerations

Several methodological limitations of the reviewed studies should be considered. Firstly, predictor and moderator analyses were primarily secondary data analyses. Studies were therefore not necessarily powered to detect the impact of baseline variables on outcomes.

Sample size recommendations are dependent on chosen statistical analysis and expected effect size of the potential predictor or moderator variable. Kraemer and Blasey (2016) recommend sample sizes of 200 to 500 participants when examining several predictors using multiple linear regression analysis. To ensure accurate and valid results for multilevel regression analysis, sample size requirements are even higher (Moineddin et al., 2007). Among reviewed studies, those completing secondary data-analyses to large-scale RCTs (e.g., Keller et al., 2000; Koscis et al., 2009b, Schramm et al., 2017) appear sufficiently powered with the exception of those conducting cross-validation analyses (Manber et al., 2008; Serbanescu et al., 2020). A priori power analysis would have been beneficial and could help ensure studies examine an appropriate number of predictor/ moderator variables for the available sample size.

Due to lack of clarity around sufficient power to detect effects, the findings should be viewed as exploratory (e.g., Serbanescu et al., 2020). Not only significant, but also non-significant findings should be interpreted with caution. For example, ethnicity was not found to be a significant predictor or moderator. However, two of the three studies examining this variable (Eisendrath et al., 2016; Manber et al., 2008) had samples of mostly White participants and therefore would have been unlikely to detect any significant associations. It is recommended that examined variables are chosen carefully and are appropriate for the available sample (Kraemer & Blasey, 2016). For significant findings, future studies are needed to ascertain the robustness of identified predictors and moderators (Kraemer et al., 2016). This may involve experimental manipulation of such potential predictors and moderators (Steketee & Chambless, 1992).

Additionally, a variable may be a significant predictor of post-intervention depression severity, but may not be a predictor of remission, relapse or drop-out. These nuances of predictor and moderator research need to be considered when conducting analysis and interpreting findings (Steketee & Chambless, 1992). Finally, the general limitations of RCTs

should be considered when drawing conclusions. RCTs often have stringent inclusion and exclusion criteria for participants, such as the included trials tending to exclude those with certain co-morbid personality disorders or high levels of suicidality (Keller et al., 2000; Kocsis et al., 2009b; Schramm et al., 2017; Probst et al., 2020). Thus, the range of presentations amongst those with persisting forms of depression may differ in clinical practice to those in RCTs, as has been found for other areas of clinical research (Humphreys & Weisner, 2000). This will impact on the generalisability of findings.

Strengths and Limitations

A strength of this review is that search terms were intentionally kept broad, with manual searching of full texts relied on to identify all relevant studies. Additionally, searching of reference lists of related articles further increases confidence in all relevant studies having been found. Given that interest in understanding what works best for whom has increased over recent years (Simon & Perlis, 2010; Kessler, 2018), it is not surprising that the majority of articles included in this review were published in the last ten years. Interestingly, only one study examined a psychological intervention for recurrent depression (Renner & Berry, 2011). The number of studies on recurrent depression was lower than anticipated. This is possibly due to many psychological interventions for recurrent depression focusing on relapse prevention, with participants in full or partial remission, and thereby not meeting the review inclusion criteria (Bockting et al., 2015).

CASP checklists were chosen for risk of bias assessment due to the simplicity in applying and appropriate checklists being available for each of the included study designs (RCT, case-control, cohort study). However, CASP checklists were designed to support the assessment of evidence. Therefore, not all questions from the CASP checklist were relevant to this review (e.g., whether study results are applicable to the local context or would improve current provision of care). CASP checklists also do not allow for in-depth assessment of

methodological approaches to predictor and moderator analyses (e.g., suitability of chosen statistical analyses) and subsequently arising sources of bias. An alternative, more detailed, quality assessment tool that could have been considered for RCTs is the Cochrane risk-of-bias tool for randomised trials (Higgins et al., 2022). Additionally, it is important to note that the quality appraisal process is reliant on adequate reporting of information within publications. Reasonable effort was made by the author to consider supplementary publications to obtain all relevant information (original trial publications, study protocols), thereby enhancing the quality appraisal process and interpretation of findings.

Meta-analysis can help increase validity and confidence in findings compared to narrative synthesis alone (Valentine et al., 2010). Additionally, meta-analysis can help detect small effect sizes by combining data from several trials (Blundell et al., 2014). Although meta-analysis can theoretically be completed with as few as two studies (Valentine et al., 2010), further meta-analytic analysis was not considered appropriate for this systematic review. For each of the variables considered for meta-analysis (age, baseline depression severity, childhood trauma) several factors contributed to lack of suitability. This includes the following: differences in experimental and control conditions, some studies utilizing the same original data source, differences in how outcomes were measured (e.g., post-intervention depression scores, depression improvement rates, % change in depression symptoms), different approaches to analyses, and lack of consistency in direction of findings (i.e., same variable may have been found to be a predictor in some studies, but a moderator or non-significant in others). As research into this area increases, meta-analyses are recommended to allow drawing of reliable conclusions.

This review was restricted to peer-reviewed publications. Therefore, the reviewed literature may be subject to publication bias, given that significant findings are more likely to be published than non-significant findings (Franco et al., 2014). Exclusion of grey literature

also means that studies not yet published may have been missed (Pappas & Williams, 2011). To reduce language bias, attempts were made to consider publications in the non-English language. Only full articles were considered for this review, and not abstracts and conference proceedings, thereby increasing the risk of up-to-date evidence being missed. A further limitation is that some full-text papers could not be obtained, therefore potential data may have been missed.

Implications for Research and Practice

Future studies into predictors and moderators for persisting forms of depression should address the discussed methodological challenges. To overcome the issue of small sample sizes that may hinder detection of small effects, combination of individual patient data from several trials is recommended. This has been done for depression more broadly, such as via patient-level meta-analyses (Bower et al., 2013) and has enabled application of more sophisticated statistical analysis (Buckman et al., 2021). Another approach included combining several promising baseline variables into a 'combined moderator' and then use statistical algorithms to help guide treatment selection. This has been successfully done for depression (Lorenzo-Luaces et al., 2017; Lorenzo-Luaces et al., 2020). In this review, only Serbanescu et al. (2020) chose such an approach. Furthermore, there is the need for a more strategic approach to analysing predictor and moderator variables within MDD research, including consistency in measuring variables and utilising similar outcome measures (Kessler et al., 2016). This would help with subsequent combination of data from different trials and enable better comparison between studies. Additionally, considered decisions should be made about which predictors and moderators are being examined, informed by the literature and in consideration of sample characteristics (Kraemer & Blasey, 2015).

At present, only baseline depression severity has been identified as a replicated and consistent predictor of treatment outcomes for persisting forms of depression. Although

individual studies found significant moderators, evidence is insufficient for utilising these moderators in clinical practice. Clinicians who work with patients with persisting forms of depression should ensure that treatment allocation is not influenced by their own biases of who they believe will benefit most from a given psychological intervention. To identify moderators that can aid treatment allocation in the UK, future moderator research should compare the various NICE guidelines recommended treatments (2022).

Conclusion

This review identified baseline depression severity as a predictor of treatment outcome for persisting forms of depression. For variables examined more frequently, findings on potential predictive and moderating effects were similar to those reported by studies on all subtypes of MDD. This suggests that predictors and moderators of treatment outcome for persisting forms of depression may be similar to those for depression overall. Although a wide range of baseline characteristics covering several domains were examined, due to limited replication of findings further research is required. In particular, moderator research for persisting forms of depression was found to be in the early stages with a very limited evidence-base. Therefore, caution should be taken with allocating patients with persisting forms of depression to psychological treatment based on baseline characteristics.

References

Amati, F., Banks, C., Greenfield, G., & Green, J. (2017). Predictors of outcomes for patients with common mental health disorders receiving psychological therapies in community settings: a systematic review. *Journal of Public Health, 40*(3), 375–387. https://doi.org/10.1093/pubmed/fdx168

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV* (Fourth edition.). American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (Fifth edition.). American Psychiatric Association.

Arnow, B. A., Manber, R., Blasey, C., Klein, D. N., Blalock, J. A., Markowitz, J. C., Rothbaum, B. O., Rush, A. J., Thase, M. E., Riso, L. P., Vivian, D., McCullough, J. P., & Keller, M. B. (2003). Therapeutic reactance as a predictor of outcome in the treatment of chronic depression. *Journal of Consulting and Clinical Psychology*, *71*(6), 1025–1035. https://doi.org/10.1037/0022-006x.71.6.1025

Arnow, B. A., Steidtmann, D., Blasey, C., Manber, R., Constantino, M. J., Klein, D. N., Markowitz, J. C., Rothbaum, B. O., Thase, M. E., Fisher, A. J., & Kocsis, J. H. (2013). The relationship between the therapeutic alliance and treatment outcome in two distinct psychotherapies for chronic depression. *Journal of Consulting and Clinical Psychology*, *81*(4), 627–638. https://doi.org/10.1037/a0031530

Assmann, N., Schramm, E., Kriston, L., Hautzinger, M., Härter, M., Schweiger, U., & Klein, J. P. (2018). Moderating effect of comorbid anxiety disorders on treatment outcome in a randomized controlled psychotherapy trial in early-onset persistently depressed outpatients. *Depression and Anxiety*, *35*(10), 1001–1008. https://doi.org/10.1002/da.22839

Bausch, P., Fangmeier, T., Meister, R., Elsaeßer, M., Kriston, L., Klein, J. P., Zobel, I., Hautzinger, M., Härter, M., & Schramm, E. (2020). The impact of childhood maltreatment on long-term outcomes in disorder-specific vs. nonspecific psychotherapy for chronic depression. *Journal of Affective Disorders*, *272*, 152–157. https://doi.org/10.1016/j.jad.2020.03.164

Bell, I. A. N., & Mellor, D. (2009). Clinical judgements: research and practice. *Australian Psychologist*, *44*(2), 112-121. https://doi.org/10.1080/00050060802550023

Berlim, M. T., & Turecki, G. (2007). What is the meaning of treatment resistant/ refractory major depression (TRD)? A systematic review of current randomized trials. *European Neuropsychopharmacology*, *17*(11), 696–707. https://doi.org/10.1016/j.euroneuro. 2007.03.009

Biesheuvel-Leliefeld, K. E. M., Kok, G. D., Bockting, C. L. H., Cuijpers, P., Hollon, S. D., van Marwijk, H. W. J., & Smit, F. (2015). Effectiveness of psychological interventions in preventing recurrence of depressive disorder: meta-analysis and meta-regression. *Journal of Affective Disorders*, *174*, 400–410. https://doi.org/10.1016/ j.jad.2014.12.016

Blundell, M. (2014). Understanding and synthesizing my numerical data. In Boland, A., Cherry, G. & Dickson, R. (eds.), *Doing a systematic review* (1st ed., pp. 99-124). SAGE Publications Ltd.

Bockting, C. L., Hollon, S. D., Jarrett, R. B., Kuyken, W., & Dobson, K. (2015). A lifetime approach to major depressive disorder: the contributions of psychological interventions in preventing relapse and recurrence. *Clinical Psychology Review*, *41*, 16–26. https://doi.org/10.1016/ j.cpr.2015.02.003

Boland, A., Cherry, G. & Dickson, R. (2014). *Doing a systematic review* (1st ed.). SAGE Publications Ltd.

Bower, P., Kontopantelis, E., Sutton, A., Kendrick, T., Richards, D. A., Gilbody, S., Knowles, S., Cuijpers, P., Andersson, G., Christensen, H., Meyer, B., Huibers, M., Smit, F., van Straten, A., Warmerdam, L., Barkham, M., Bilich, L., Lovell, K., & Liu, E. T.-H. . (2013). Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. *BMJ, 346*, f540–f540. https://doi.org/10.1136/bmj.f540

Brown, S., Rittenbach, K., Cheung, S., McKean, G., MacMaster, F. P., & Clement, F. (2019). Current and common definitions of treatment-resistant depression: findings from a systematic review and qualitative interviews. *The Canadian Journal of Psychiatry*, *64*(6), 380–387. https://doi.org/10.1177/0706743719828965

Buckman, J. E. J., Cohen, Z. D., O'Driscoll, C., Fried, E. I., Saunders, R., Ambler, G., DeRubeis, R. J., Gilbody, S., Hollon, S. D., Kendrick, T., Watkins, E., Eley, T. C., Peel, A. J., Rayner, C., Kessler, D., Wiles, N., Lewis, G., & Pilling, S. (2021). Predicting prognosis for adults with depression using individual symptom data: a comparison of modelling approaches. *Psychological Medicine*, *53*(2) 1–11. https://doi.org/ 10.1017/s00332 91721001616

Button, K. S., Turner, N., Campbell, J., Kessler, D., Kuyken, W., Lewis, G., Peters, T. J., Thomas, L., & Wiles, N. (2015). Moderators of response to cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care. *Journal of Affective Disorders*, *174*, 272–280. https://doi.org/ 10.1016/j.jad.2014.11.057

Childhood Trauma Meta-Analysis Study Group. (2022). Treatment efficacy and effectiveness in adults with major depressive disorder and childhood trauma history: a systematic review and meta-analysis. *The Lancet Psychiatry*, S2215-0366. https://doi.org/ 10.1016/s2215-0366(22)00227-9

Cladder-Micus, M. B., Speckens, A. E. M., Vrijsen, J. N., T Donders, A. R., Becker, E. S., & Spijker, J. (2018). Mindfulness-based cognitive therapy for patients with chronic, treatment-resistant depression: a pragmatic randomized controlled trial. *Depression and Anxiety*, *35*(10), 914–924. https://doi.org/10.1002/da.22788

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/00131644 6002000104

Critical Appraisal Skills Programme. (2022). *CASP checklist*. https://casp-uk.net/casp-tools-checklists/

Crown, W. H., Finkelstein, S., Berndt, E. R., Ling, D., Poret, A. W., Rush, A. J., & Russell, J. M. (2002). The impact of treatment-resistant depression on health care utilization and costs. *The Journal of Clinical Psychiatry*, *63*(11), 963–971. https://doi.org/10.4088/ jcp.v63n1102

Cuijpers, P., Huibers, M. J. H., & Furukawa, T. A. (2017). The need for research on treatments of chronic depression. *JAMA Psychiatry*, *74*(3), 242. https://doi.org/10.1001/ jamapsychiatry. 2016.4070

Cuijpers, P., Karyotaki, E., Eckshtain, D., Ng, M. Y., Corteselli, K. A., Noma, H., Quero, S., & Weisz, J. R. (2020). Psychotherapy for depression across different age groups. *JAMA Psychiatry*, *77*(7), 694-702. https://doi.org/10.1001/jamapsychiatry.2020.0164

Cuijpers, P., van Straten, A., Schuurmans, J., van Oppen, P., Hollon, S. D., & Andersson, G. (2010). Psychotherapy for chronic major depression and dysthymia: a meta-analysis. *Clinical Psychology Review*, *30*(1), 51–62. https://doi.org/10.1016/ j.cpr.2009.09.003

Cuijpers, P., Weitz, E., Twisk, J., Kuehner, C., Cristea, I., David, D., DeRubeis, R. J., Dimidjian, S., Dunlop, B. W., Faramarzi, M., Hegerl, U., Jarrett, R. B., Kennedy, S. H.,

Kheirkhah, F., Mergl, R., Miranda, J., Mohr, D. C., Segal, Z. V., Siddique, J., & Simons, A. D. (2014). Gender as predictor and moderator of outcome in cognitive behavior therapy and pharmacotherapy for adult depression: an "individual patient data" meta-analysis. *Depression and Anxiety*, *31*(11), 941–951. https://doi.org/10.1002/da.22328

De Carlo, V., Calati, R., & Serretti, A. (2016). Socio-demographic and clinical predictors of non-response/non-remission in treatment resistant depressed patients: a systematic review. *Psychiatry Research*, *240*, 421–430. https://doi.org/10.1016/j.psychres.2016.04.034

Demyttenaere, K., & Van Duppen, Z. (2018). The impact of (the concept of) treatment-resistant depression: an opinion review. *International Journal of Neuropsychopharmacology*, *22*(2), 85–92. https://doi.org/10.1093/ijnp/pyy052

Denton, W. H., Carmody, T. J., A. John Rush, Thase, M. E., Trivedi, M. H., Arnow, B. A., Klein, K., & Keller, M. (2010). Dyadic discord at baseline is associated with lack of remission in the acute treatment of chronic depression. *Psychological Medicine*, *40*(3), 415–424. https://doi.org/10.1017/s0033291709990535

Driessen, E., Cuijpers, P., Hollon, S. D., & Dekker, J. J. M. (2010). Does pretreatment severity moderate the efficacy of psychological treatment of adult outpatient depression? A meta-analysis. *Journal of Consulting and Clinical Psychology*, *78*(5), 668–680. https://doi.org/10.1037/a0020570

Eisendrath, S. J., Gillung, E., Delucchi, K. L., Segal, Z. V., Nelson, J. C., McInnes, L. A., Mathalon, D. H., & Feldman, M. D. (2016). A randomized controlled trial of mindfulness-based cognitive therapy for treatment-resistant depression. *Psychotherapy and Psychosomatics*, *85*(2), 99–110. https://doi.org/10.1159/000442260

Eriksen, M. B., & Frandsen, T. F. (2018). The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. *Journal of the Medical Library Association*, *106*(4), 420–431. https://doi.org/10.5195/jmla.2018.345

Erkens, N., Schramm, E., Kriston, L., Hautzinger, M., Härter, M., Schweiger, U., & Klein, J. P. (2018). Association of comorbid personality disorders with clinical characteristics and outcome in a randomized controlled trial comparing two psychotherapies for early-onset persistent depressive disorder. *Journal of Affective Disorders*, *229*, 262–268. https://doi.org/10.1016/j.jad.2017.12.091

Fonagy, P., Rost, F., Carlyle, J., McPherson, S., Thomas, R., Pasco Fearon, R. M., Goldberg, D., & Taylor, D. (2015). Pragmatic randomized controlled trial of long-term psychoanalytic psychotherapy for treatment-resistant depression: the Tavistock adult depression study (TADS). *World Psychiatry*, *14*(3), 312–321. https://doi.org/10.1002/wps.20267

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: unlocking the file drawer. *Science*, *345*(6203), 1502–1505. https://doi.org/10.1126/science.1255484

Gopinath, S., Katon, W. J., Russo, J. E., & Ludman, E. J. (2007). Clinical factors associated with relapse in primary care patients with chronic or recurrent depression. *Journal of Affective Disorders*, *101*(1-3), 57–63. https://doi.org/10.1016/j.jad.2006.10.023

Greden, J. F. (2001). The burden of recurrent depression: causes, consequences, and future prospects. *The Journal of Clinical Psychiatry*, *62 Suppl 22*, 5–9. https://pubmed.ncbi.nlm.nih.gov/11599650/

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for

treatment failure. *Journal of Clinical Psychology*, *61*(2), 155–163. https://doi.org/10.1002/jclp.20108

Hardeveld, F., Spijker, J., De Graaf, R., Nolen, W. A., & Beekman, A. T. F. (2009). Prevalence and predictors of recurrence of major depressive disorder in the adult population. *Acta Psychiatrica Scandinavica*, *122*(3), 184–191. https://doi.org/10.1111/j.16000447.2009. 01519.x

Higgins, J., Savović, J., Page, M., Elbers, R., Sterne, J. (2022). Assessing risk of bias in a randomized trial. In Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (Eds). *Cochrane handbook for systematic reviews of interventions* (Version 6.3). Cochrane. www.training.cochrane.org/handbook.

Humphreys, K., & Weisner, C. (2000). Use of exclusion criteria in selecting research subjects and its effect on the generalizability of alcohol treatment outcome studies. *American Journal of Psychiatry*, *157*(4), 588–594. https://doi.org/10.1176/appi.ajp.157.4.588

ICD-10 World Health Organization(WHO). (1993). *The ICD-10 classification of mental and behavioural disorders*. World Health Organization.

Ijaz, S., Davies, P., Williams, C. J., Kessler, D., Lewis, G., & Wiles, N. (2018). Psychological therapies for treatment-resistant depression in adults. *Cochrane Database of Systematic Reviews*, *5*(CD010558). https://doi.org/10.1002/14651858.cd010558.pub2

Kaiser, T., Volkmann, C., Volkmann, A., Karyotaki, E., Cuijpers, P., & Brakemeier, E.-L. (2022). Heterogeneity of treatment effects in trials on psychotherapy of depression. *Clinical Psychology: Science and Practice*, *29*(3), 294–303. https://doi.org/10.1037/cps0000079

Keller, M. B., McCullough, J. P., Klein, D. N., Arnow, B., Dunner, D. L., Gelenberg, A. J., Markowitz, J. C., Nemeroff, C. B., Russell, J. M., Thase, M. E., Trivedi, M. H., Blalock, J. A., Borian, F. E., Jody, D. N., DeBattista, C., Koran, L. M., Schatzberg, A. F.,

Fawcett, J., Hirschfeld, R. M. A., & Keitner, G. (2000). A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *New England Journal of Medicine*, *342*(20), 1462–1470. https://doi.org/10.1056/nejm200005183422001

Kessler, R. C. (2018). The potential of predictive analytics to provide clinical decision support in depression treatment planning. *Current Opinion in Psychiatry*, *31*(1), 32–39. https://doi.org/10.1097/yco.0000000000000377

Kessler, R. C., van Loo, H. M., Wardenaar, K. J., Bossarte, R. M., Brenner, L. A., Ebert, D. D., de Jonge, P., Nierenberg, A. A., Rosellini, A. J., Sampson, N. A., Schoevers, R. A., Wilcox, M. A., & Zaslavsky, A. M. (2016). Using patient self-reports to study heterogeneity of treatment effects in major depressive disorder. *Epidemiology and Psychiatric Sciences*, *26*(1), 22–36. https://doi.org/10.1017/s2045796016000020

Kiosses, D. N., Leon, A. C., & Areán, P. A. (2011). Psychosocial interventions for late-life major depression: evidence-based treatments, predictors of treatment outcomes, and moderators of treatment effects. *Psychiatric Clinics of North America*, *34*(2), 377–401. https://doi.org/10.1016/j.psc.2011.03.001

Klein, J. P., Erkens, N., Schweiger, U., Kriston, L., Bausch, P., Zobel, I., Hautzinger, M., Schoepf, D., Serbanescu, I., Bailer, J., Backenstrass, M., Wambach, K., Walter, H., Härter, M., & Schramm, E. (2018). Does childhood maltreatment moderate the effect of the cognitive behavioral analysis system of psychotherapy versus supportive psychotherapy in persistent depressive disorder? *Psychotherapy and Psychosomatics*, *87*(1), 46–48. https://doi.org/10.1159/000484412

Kocsis, J. H., Gelenberg, A., Rothbaum, B. O., Klein, D. N., Trivedi, M., Manber, R., Keller, M. B., Leon, A. C., Wisniewski, S. R., Arnow, B. A., Markowitz, J. C., & Thase, M. (2009b). Cognitive behavioral analysis system of psychotherapy and brief supportive

psychotherapy for augmentation of antidepressant nonresponse in chronic depression. The REVAMP trial augmentation with CBASP and BSP for depression. *Archives of General Psychiatry*, *66*(11), 1178-1188. https://doi.org/10.1001/archgenpsychiatry.2009.144

Kocsis, J. H., Leon, A. C., Markowitz, J. C., Manber, R., Arnow, B., Klein, D. N., & Thase, M. E. (2009a). Patient preference as a moderator of outcome for chronic forms of major depressive disorder treated with nefazodone, cognitive behavioral analysis system of psychotherapy, or their combination. *The Journal of Clinical Psychiatry*, *70*(3), 354–361. https://doi.org/10.4088/jcp.08m04371

Kraemer, H. C. (2016). Messages for clinicians: moderators and mediators of treatment outcome in randomized clinical trials. *American Journal of Psychiatry*, *173*(7), 672–679. https://doi.org/10.1176/appi.ajp.2016.15101333

Kraemer, H.C., & Blasey, C. (2016). Linear Regression Analysis. In Kraemer, H.C., & Blasey, C. (Eds), *How many subjects?: statistical power analysis in research*. (2nd ed.). SAGE Publications Ltd. https://doi.org/10.4135/9781483398761

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59*(10), 877-883 https://doi.org/10.1001/archpsyc.59.10.877

Landis J, R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Li, J.-M., Zhang, Y., Su, W.-J., Liu, L.-L., Gong, H., Peng, W., & Jiang, C.-L. (2018). Cognitive behavioral therapy for treatment-resistant depression: a systematic review and meta-analysis. *Psychiatry Research*, *268*, 243–250. https://doi.org/10.1016/j.psychres.2018.07.020

Lopez, M. A., & Basco, M. A. (2015). Effectiveness of cognitive behavioral therapy in public mental health: comparison to treatment as usual for treatment-resistant depression. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(1), 87–98. https://doi.org/10.1007/s10488-014-0546-4

Lorenzo-Luaces, L., DeRubeis, R. J., van Straten, A., & Tiemens, B. (2017). A prognostic index (PI) as a moderator of outcomes in the treatment of depression: a proof of concept combining multiple variables to inform risk-stratified stepped care models. *Journal of Affective Disorders*, *213*, 78–85. https://doi.org/10.1016/j.jad.2017.02.010

Lorenzo-Luaces, L., Peipert, A., De Jesús Romero, R., Rutter, L. A., & Rodriguez-Quintana, N. (2020). Personalized medicine and cognitive behavioral therapies for depression: small effects, big problems, and bigger data. *International Journal of Cognitive Therapy, 14*, 59-85. https://doi.org/10.1007/s41811-020-00094-3

Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies. *Archives of General Psychiatry*, *32*(8), 995. https://doi.org/10.1001/archpsyc.1975.01760260059004

Manber, R., Kraemer, H. C., Arnow, B. A., Trivedi, M. H., Rush, A. J., Thase, M. E., Rothbaum, B. O., Klein, D. N., Kocsis, J. H., Gelenberg, A. J., & Keller, M. E. (2008). Faster remission of chronic depression with combined psychotherapy and medication than with each therapy alone. *Journal of Consulting and Clinical Psychology*, *76*(3), 459–467. https://doi.org/10.1037/0022-006x.76.3.459

McAllister-Williams, R. H., Arango, C., Blier, P., Demyttenaere, K., Falkai, P., Gorwood, P., Hopwood, M., Javed, A., Kasper, S., Malhi, G. S., Soares, J. C., Vieta, E., Young, A. H., Papadopoulos, A., & Rush, A. J. (2020). The identification, assessment, and management of difficult-to-treat depression: an international consensus

statement. *Journal of Affective Disorders*, *267*, 264–282. https://doi.org/10.1016/j.jad.2020.02.023

McCullough, J. P., Jr. (2000). *Treatment for chronic depression: cognitive behavioral analysis system of psychotherapy (CBASP).* Guilford Press.

McLaughlin, K. A. (2011). The public health impact of major depression: a call for interdisciplinary prevention efforts. *Prevention Science*, *12*(4), 361–371. https://doi.org/10.1007/s11121-011-0231-8

McPherson, S., & Senra, H. (2022). Psychological treatments for persistent depression: a systematic review and meta-analysis of quality of life and functioning outcomes. *Psychotherapy, 59*(3), 447-459. https://doi.org/10.1037/pst0000448

McPherson, S., Cairns, P., Carlyle, J., Shapiro, D. A., Richardson, P., & Taylor, D. (2005). The effectiveness of psychological treatments for treatment-resistant depression: a systematic review. *Acta Psychiatrica Scandinavica*, *111*(5), 331–340. https://doi.org/10.1111/j.1600-0447.2004.00498.x

Michalak, J., Schultze, M., Heidenreich, T., & Schramm, E. (2015). A randomized controlled trial on the efficacy of mindfulness-based cognitive therapy and a group version of cognitive behavioral analysis system of psychotherapy for chronically depressed patients. *Journal of Consulting and Clinical Psychology, 83*(5), 951–963. https://doi.org/10.1037/ccp0000042

Moineddin, R., Matheson, F. I., & Glazier, R. H. (2007). A simulation study of sample size for multilevel logistic regression models. *BMC Medical Research Methodology*, *7*(1). https://doi.org/10.1186/1471-2288-7-34

Mueller, T. I., Leon, A. C., Keller, M. B., Solomon, D. A., Endicott, J., Coryell, W., Warshaw, M., & Maser, J. D. (1999). Recurrence after recovery from major depressive disorder

during 15 years of observational follow-up. *The American Journal of Psychiatry*, *156*(7), 1000–1006. https://doi.org/10.1176/ajp.156.7.1000

Murphy, J. A., & Byrne, G. J. (2012). Prevalence and correlates of the proposed DSM-5 diagnosis of chronic depressive disorder. *Journal of Affective Disorders*, *139*(2), 172–180. https://doi.org/10.1016/j.jad.2012.01.033

Nanni, V., Uher, R., & Danese, A. (2012). Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: a meta-analysis. *American Journal of Psychiatry*, *169*(2), 141–151. https://doi.org/10.1176/appi.ajp.2011.11020335

National Institute for Health and Care Excellence (NICE, 2022). *Depression in adults: recognition and management*. [NICE Guideline No. 222]. https://www.nice.org.uk/guidance/ng222

Nelson, J., Klumparendt, A., Doebler, P., & Ehring, T. (2017). Childhood maltreatment and characteristics of adult depression: meta-analysis. *British Journal of Psychiatry*, *210*(2), 96–104. https://doi.org/10.1192/bjp.bp.115.180752

Nemeroff, C. B., Heim, C. M., Thase, M. E., Klein, D. N., Rush, A. J., Schatzberg, A. F., Ninan, P. T., McCullough, J. P., Weiss, P. M., Dunner, D. L., Rothbaum, B. O., Kornstein, S., Keitner, G., & Keller, M. B. (2003). Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. *Proceedings of the National Academy of Sciences*, *100*(24), 14293–14296. https://doi.org/10.1073/pnas. 2336126100

Nilsen, T. S., Eisemann, M., & Kvernmo, S. (2012). Predictors and moderators of outcome in child and adolescent anxiety and depression: a systematic review of psychological treatment studies. *European Child & Adolescent Psychiatry*, *22*(2), 69–87. https://doi.org/10.1007/ s00787-012-0316-3

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & McGuinness, L. A. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *PLOS Medicine*, *18*(3), e1003583. https://doi.org/10.1371/ journal.pmed.1003583

Palpacuer, C., Gallet, L., Drapier, D., Reymann, J.-M., Falissard, B., & Naudet, F. (2017). Specific and non-specific effects of psychotherapeutic interventions for depression: results from a meta-analysis of 84 studies. *Journal of Psychiatric Research*, *87*, 95–104. https://doi.org/10.1016/j.jpsychires.2016.12.015

Papakostas, G. I., & Fava, M. (2022). Predictors, moderators, and mediators (correlates) of treatment outcome in major depressive disorder. *Dialogues in Clinical Neuroscience, 10*(4), 439-451. https://doi.org/10.31887/DCNS.2008.10.4/gipapakostas

Pappas, C., & Williams, I. (2011). Grey literature: its emerging importance. *Journal of Hospital Librarianship*, *11*(3), 228–234. https://doi.org/10.1080/15323269.2011.587100

Potijk, M. R., Aan het Rot, M., Parlevliet, F. M., Schoevers, R. A., & Eldering, M. J. (2020). Group cognitive behavioural analysis system of psychotherapy (CBASP) for persistently depressed outpatients: a retrospective chart review. *British Journal of Clinical Psychology*, *59*(4), 552–564. https://doi.org/10.1111/bjc.12266

Probst, T., Schramm, E., Heidenreich, T., Klein, J., & Michalak, J. (2020). Patients' interpersonal problems as moderators of depression outcomes in a randomized controlled trial comparing mindfulness-based cognitive therapy and a group version of the cognitive-behavioral analysis system of psychotherapy in chronic depression. *Journal of Clinical Psychology*, *76*(7), 1241-1254. https://doi.org/10.1002/jclp.22931

Renner, W., & Berry, J. W. (2011). The ineffectiveness of group interventions for female Turkish migrants with recurrent depression. *Social Behavior and Personality: An International Journal*, *39*(9), 1217–1234. https://doi.org/10.2224/sbp.2011.39.9.1217

Rost, F., Luyten, P., Fearon, P., & Fonagy, P. (2019). Personality and outcome in individuals with treatment-resistant depression—exploring differential treatment effects in the Tavistock adult depression study (TADS). *Journal of Consulting and Clinical Psychology*, *87*(5), 433–445. https://doi.org/10.1037/ccp0000391

Schramm, E., Klein, D. N., Elsaesser, M., Furukawa, T. A., & Domschke, K. (2020). Review of dysthymia and persistent depressive disorder: history, correlates, and clinical implications. *The Lancet Psychiatry*, *7*(9), 801–812. https://doi.org/10.1016/S2215-0366(20)30099-7

Schramm, E., Kriston, L., Elsaesser, M., Fangmeier, T., Meister, R., Bausch, P., Zobel, I., Bailer, J., Wambach, K., Backenstrass, M., Klein, J. P., Schoepf, D., Gumz, A., Loewe, B., Walter, H., Wolf, M., Domschke, K., Berger, M., Hautzinger, M., & Härter, M. (2019). Two-year follow-up after treatment with the cognitive behavioral analysis system of psychotherapy versus supportive psychotherapy for early-onset chronic depression. *Psychotherapy and Psychosomatics*, *88*(3), 154-164. https://doi-org.sheffield.idm.oclc.org/10.1159/000500189

Schramm, E., Kriston, L., Zobel, I., Bailer, J., Wambach, K., Backenstrass, M., Klein, J. P., Schoepf, D., Schnell, K., Gumz, A., Bausch, P., Fangmeier, T., Meister, R., Berger, M., Hautzinger, M., & Härter, M. (2017). Effect of disorder-specific vs nonspecific psychotherapy for chronic depression. *JAMA Psychiatry*, *74*(3), 233-242. https://doi.org/10.1001/ jamapsychiatry.2016. 3880

Serbanescu, I., Walter, H., Schnell, K., Kessler, H., Weber, B., Drost, S., Groß, M., Neudeck, P., Klein, J. P., Assmann, N., Zobel, I., Backenstrass, M., Hautzinger, M., Meister, R.,

Härter, M., Schramm, E., & Schoepf, D. (2020). Combining baseline characteristics to disentangle response differences to disorder-specific versus supportive psychotherapy in patients with persistent depressive disorder. *Behaviour Research and Therapy*, *124*, 103512. https://doi.org/10.1016/j.brat.2019.103512

Shankman, S. A., Campbell, M. L., Klein, D. N., Leon, A. C., Arnow, B. A., Manber, R., Keller, M. B., Markowitz, J. C., Rothbaum, B. O., Thase, M. E., & Kocsis, J. H. (2013). Dysfunctional attitudes as a moderator of pharmacotherapy and psychotherapy for chronic depression. *Journal of Psychiatric Research*, *47*(1), 113–121. https://doi.org/10.1016/j.jpsychires. 2012.09.018

Sharpley, C. F., & Bitsika, V. (2011). Four potential criteria for deciding when to use antidepressants or psychotherapy for unipolar depression: a literature review. *International Journal of Psychiatry in Clinical Practice*, *15*(1), 2–11. https://doi.org/10.3109/13651501.2010.527008

Simon, G. E., & Perlis, R. H. (2010). Personalized medicine for depression: can we match patients with treatments? *American Journal of Psychiatry*, *167*(12), 1445–1455. https://doi.org/10.1176/appi.ajp.2010.09111680

Smith, D. J., Nicholl, B. I., Cullen, B., Martin, D., Ul-Haq, Z., Evans, J., Gill, J. M. R., Roberts, B., Gallacher, J., Mackay, D., Hotopf, M., Deary, I., Craddock, N., & Pell, J. P. (2013). Prevalence and characteristics of probable major depression and bipolar disorder within uk biobank: cross-sectional study of 172,751 participants. *PLoS ONE*, *8*(11), e75362. https://doi.org/10.1371/journal.pone.0075362

Stangier, U., Frick, A., Thinnes, I., Arens, E. A., & Hofmann, S. G. (2021). Metta-based therapy for chronic depression: a wait list control trial. *Mindfulness*, *12*(12), 2929–2942. https://doi.org/10.1007/s12671-021-01753-y

Steidtmann, D., Manber, R., Arnow, B. A., Klein, D. N., Markowitz, J. C., Rothbaum, B. O., Thase, M. E., & Kocsis, J. H. (2012). Patient treatment preference as a predictor of response and attrition in treatment for chronic depression. *Depression and Anxiety*, *29*(10), 896–905. https://doi.org/10.1002/da.21977

Steketee, G. (1992). Methodological issues in prediction of treatment outcome. *Clinical Psychology Review*, *12*(4), 387–400. https://doi.org/10.1016/0272-7358(92)90123-p

Tanguay-Sela, M., Rollins, C., Perez, T., Qiang, V., Golden, G., Tunteng, J.-F., Perlman, K., Simard, J., Benrimoh, D., & Margolese, H. C. (2022). A systematic meta-review of patient-level predictors of psychological therapy outcome in major depressive disorder. *Journal of Affective Disorders*, *317*, 307–318. https://doi.org/10.1016/j.jad.2022.08.041

Taubner, S., Kessler, H., Buchheim, A., Kächele, H., & Staun, L. (2011). The role of mentalization in the psychoanalytic treatment of chronic depression. *Psychiatry: Interpersonal and Biological Processes*, *74*(1), 49–57. https://doi.org/10.1521/psyc.2011.74.1.49

Tenny, S., Kerndt, C. C., & Hoffman, M. R. (2020). *Case control studies*. StatPearls Publishing. https://pubmed.ncbi.nlm.nih.gov/28846237/

Thomas, L., Kessler, D., Campbell, J., Morrison, J., Peters, T. J., Williams, C., Lewis, G., & Wiles, N. (2013). Prevalence of treatment-resistant depression in primary care: cross-sectional data. *British Journal of General Practice*, *63*(617), e852–e858. https://doi.org/ 10.3399/bjgp13x675430

Torpey, D. C., & Klein, D. N. (2008). Chronic depression: update on classification and treatment. *Current Psychiatry Reports*, *10*(6), 458–464. https://doi.org/10.1007/s11920-008-0074-6

Tunvirachaisakul, C., Gould, R. L., Coulson, M. C., Ward, E. V., Reynolds, G., Gathercole, R. L., Grocott, H., Supasitthumrong, T., Tunvirachaisakul, A., Kimona, K., & Howard, R. J. (2018). Predictors of treatment outcome in depression in later life: a systematic review and meta-analysis. *Journal of Affective Disorders*, *227*, 164–182. https://doi.org/10.1016/ j.jad.2017.10.008

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? *Journal of Educational and Behavioral Statistics*, *35*(2), 215–247. https://doi.org/10.3102/ 1076998609346961

van Bronswijk, S., Moopen, N., Beijers, L., Ruhe, H. G., & Peeters, F. (2018). Effectiveness of psychotherapy for treatment-resistant depression: a meta-analysis and meta-regression. *Psychological Medicine*, *49*(3), 366–379. https://doi.org/10.1017/ s00332917 1800199x

Weitz, E. S., Hollon, S. D., Twisk, J., van Straten, A., Huibers, M. J. H., David, D., DeRubeis, R. J., Dimidjian, S., Dunlop, B. W., Cristea, I. A., Faramarzi, M., Hegerl, U., Jarrett, R. B., Kheirkhah, F., Kennedy, S. H., Mergl, R., Miranda, J., Mohr, D. C., Rush, A. J., & Segal, Z. V. (2015). Baseline depression severity as moderator of depression outcomes between cognitive behavioral therapy vs pharmacotherapy. *JAMA Psychiatry*, *72*(11), 1102-1109. https://doi.org/10.1001/jamapsychiatry.2015.1516

Wiles, N., Thomas, L., Abel, A., Ridgway, N., Turner, N., Campbell, J., Garland, A., Hollinghurst, S., Jerrom, B., Kessler, D., Kuyken, W., Morrison, J., Turner, K., Williams, C., Peters, T., & Lewis, G. (2013). Cognitive behavioural therapy as an adjunct to pharmacotherapy for primary care-based patients with treatment resistant depression: results of the CoBalT randomised controlled trial. *The Lancet*, *381*(9864), 375–384. https://doi.org/10.1016/s0140-6736(12)61552-9

Appendices

Appendix A

The Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA)2020 Item Checklist (Page et al., 2021)

PRISMA 2020 Main Checklist

Topic	No.	Item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	p.2, p. 9
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	pp. 5-9
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	pp. 9
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	pp. 11-13
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	p. 10
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	pp. 10-11
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	pp. 14-15

Topic	No.	Item	Location where item is reported
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	p. 16
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	p. 16
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	p. 16
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	p. 16
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	not applicable
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item 5)).	not applicable
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	not applicable
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	p. 16

Topic	No.	Item	Location where item is reported
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	p. 16, p. 43
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	not applicable
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	not applicable
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	not applicable
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	not applicable
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	pp. 14-15
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	p. 14
Study characteristics	17	Cite each included study and present its characteristics.	pp. 17-23
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	p. 24
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	pp. 25-29
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	p. 24

Topic	No.	Item	Location where item is reported
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	not applicable
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	not applicable
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	not applicable
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	p. 24
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	p. 42
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	pp. 36-40
	23b	Discuss any limitations of the evidence included in the review.	pp. 40-42
	23c	Discuss any limitations of the review processes used.	pp. 42-44
	23d	Discuss implications of the results for practice, policy, and future research.	pp. 44-45
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	p. 10
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	p. 10

Topic	No.	Item	Location where item is reported
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	not applicable
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	not applicable
Competing interests	26	Declare any competing interests of review authors.	not applicable
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	not applicable

PRISMA Abstract Checklist

Topic	No.	Item	Reported?
TITLE			
Title	1	Identify the report as a systematic review.	Yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.	Yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	No
Information sources	4	Specify the information sources (e.g. databases, registers) used to identify studies and the date when each was last searched.	Yes
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.	Yes
Synthesis of results	6	Specify the methods used to present and synthesize results.	Yes
RESULTS			

Topic	No.	Item	Reported?
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.	Yes
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e. which group is favoured).	Yes
DISCUSSION			
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g. study risk of bias, inconsistency and imprecision).	Yes
Interpretation	10	Provide a general interpretation of the results and important implications.	Yes
OTHER			
Funding	11	Specify the primary source of funding for the review.	No
Registration	12	Provide the register name and registration number.	Yes

Appendix B

Critical Appraisal Skills Programme (CASP, 2022) Checklist - Case-Control Studies

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?	Yes ☐ Can't Tell ☐ No ☐	HINT: An issue can be 'focused' In terms of • the population studied • Whether the study tried to detect a beneficial or harmful effect • the risk factors studied

Comments:

2. Did the authors use an appropriate method to answer their question?	Yes ☐ Can't Tell ☐ No ☐	HINT: Consider • Is a case control study an appropriate way of answering the question under the circumstances • Did it address the study question

Comments:

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes ☐
Can't Tell ☐
No ☐

Comments:

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes ☐
Can't Tell ☐
No ☐

Comments:

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

5. Was the exposure accurately measured to minimise bias?

Yes ☐
Can't Tell ☐
No ☐

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

Comments:

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List:

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes ☐
Can't Tell ☐
No ☐

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments:

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments:

9. Do you believe the results?

Yes ☐
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments:

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☐
Can't Tell ☐
No ☐

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments:

11. Do the results of this study fit with other available evidence?

Yes ☐
Can't Tell ☐
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments:

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.

Appendix C

Critical Appraisal Skills Programme (CASP, 2022) Checklist - Cohort Studies

Section A: Are the results of the study valid?

1. Did the study address a clearly focused issue?

Yes ☐
Can't Tell ☐
No ☐

HINT: A question can be 'focused' in terms of
- the population studied
- the risk factors studied
- is it clear whether the study tried to detect a beneficial or harmful effect
- the outcomes considered

Comments:

2. Was the cohort recruited in an acceptable way?

Yes ☐
Can't Tell ☐
No ☐

HINT: Look for selection bias which might compromise the generalisability of the findings:
- was the cohort representative of a defined population
- was there something special about the cohort
- was everybody included who should have been

Comments:

Is it worth continuing?

3. Was the exposure accurately measured to minimise bias?

Yes ☐
Can't Tell ☐
No ☐

HINT: Look for measurement or classification bias:

- did they use subjective or objective measurements
- do the measurements truly reflect what you want them to (have they been validated)
- were all the subjects classified into exposure groups using the same procedure

Comments:	

4. Was the outcome accurately measured to minimise bias?

Yes ☐
Can't Tell ☐
No ☐

HINT: Look for measurement or classification bias:

- did they use subjective or objective measurements
- do the measurements truly reflect what you want them to (have they been validated)
- has a reliable system been established for detecting all the cases (for measuring disease occurrence)
- were the measurement methods similar in the different groups
- were the subjects and/or the outcome assessor blinded to exposure (does this matter)

Comments:	

5. (a) Have the authors identified all important confounding factors?

Yes ☐
Can't Tell ☐
No ☐

HINT:
- list the ones you think might be important, and ones the author missed

Comments:

5. (b) Have they taken account of the confounding factors in the design and/or analysis?

Yes ☐
Can't Tell ☐
No ☐

HINT:
- look for restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments:

6. (a) Was the follow up of subjects complete enough?

Yes ☐
Can't Tell ☐
No ☐

HINT: Consider
- the good or bad effects should have had long enough to reveal themselves
- the persons that are lost to follow-up may have different outcomes than those available for assessment
- in an open or dynamic cohort, was there anything special about the outcome of the people leaving, or the exposure of the people entering the cohort

6. (b) Was the follow up of subjects long enough?

Yes ☐
Can't Tell ☐
No ☐

Comments:	

Section B: What are the results?

7. What are the results of this study?

HINT: Consider
- what are the bottom line results
- have they reported the rate or the proportion between the exposed/unexposed, the ratio/rate difference
- how strong is the association between exposure and outcome (RR)
- what is the absolute risk reduction (ARR)

Comments:	

8. How precise are the results?

HINT:
- look for the range of the confidence intervals, if given

Comments:	

9. Do you believe the results?

Yes ☐
Can't Tell ☐
No ☐

HINT: Consider
- big effect is hard to ignore
- can it be due to bias, chance or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- Bradford Hills criteria (e.g. time sequence, dose-response gradient, biological plausibility, consistency)

Comments:

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☐
Can't Tell ☐
No ☐

HINT: Consider whether
- a cohort study was the appropriate method to answer this question
- the subjects covered in this study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- you can quantify the local benefits and harms

Comments:

11. Do the results of this study fit with other available evidence?

Yes ☐
Can't Tell ☐
No ☐

Comments:

12. What are the implications of this study for practice?

Yes ☐

Can't Tell ☐

No ☐

HINT: Consider

- one observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making
- for certain questions, observational studies provide the only evidence
- recommendations from observational studies are always stronger when supported by other evidence

Comments:	

Appendix D

Critical Appraisal Skills Programme (CASP, 2022) Checklist - Randomised Controlled Trial

Section A: Is the basic study design valid for a randomised controlled trial?			
1. Did the study address a clearly focused research question? *CONSIDER:* *Was the study designed to assess the outcomes of an intervention?* *Is the research question 'focused' in terms of:* • *Population studied* • *Intervention given* • *Comparator chosen* • *Outcomes measured?*	Yes ☐	No ☐	Can't tell ☐
2. Was the assignment of participants to interventions randomised? *CONSIDER:* • *How was randomisation carried out? Was the method appropriate?* • *Was randomisation sufficient to eliminate systematic bias?* • *Was the allocation sequence concealed from investigators and participants?*	Yes ☐	No ☐	Can't tell ☐
3. Were all participants who entered the study accounted for at its conclusion? *CONSIDER:* • *Were losses to follow-up and exclusions after randomisation accounted for?* • *Were participants analysed in the study groups to which they were randomised (intention-to-treat analysis)?* • *Was the study stopped early? If so, what was the reason?*	Yes ☐	No ☐	Can't tell ☐

Section B: Was the study methodologically sound?			
4. • **Were the participants 'blind' to intervention they were given?** • **Were the investigators 'blind' to the intervention they were giving to participants?** • **Were the people assessing/analysing outcome/s 'blinded'?**	Yes ☐ ☐ ☐	No ☐ ☐ ☐	Can't tell ☐ ☐ ☐
5. Were the study groups similar at the start of the randomised controlled trial? *CONSIDER:* • *Were the baseline characteristics of each study group (e.g. age, sex, socio-economic group) clearly set out?* • *Were there any differences between the study groups that could affect the outcome/s?*	Yes ☐	No ☐	Can't tell ☐

6. Apart from the experimental intervention, did each study group receive the same level of care (that is, were they treated equally)? *CONSIDER:* • *Was there a clearly defined study protocol?* • *If any additional interventions were given (e.g. tests or treatments), were they similar between the study groups?* • *Were the follow-up intervals the same for each study group?*	Yes ☐	No ☐	Can't tell ☐

Section C: What are the results?

7. Were the effects of intervention reported comprehensively? *CONSIDER:* • *Was a power calculation undertaken?* • *What outcomes were measured, and were they clearly specified?* • *How were the results expressed? For binary outcomes, were relative and absolute effects reported?* • *Were the results reported for each outcome in each study group at each follow-up interval?* • *Was there any missing or incomplete data?* • *Was there differential drop-out between the study groups that could affect the results?* • *Were potential sources of bias identified?* • *Which statistical tests were used?* • *Were p values reported?*	Yes ☐	No ☐	Can't tell ☐
8. Was the precision of the estimate of the intervention or treatment effect reported? *CONSIDER:* • *Were confidence intervals (CIs) reported?*	Yes ☐	No ☐	Can't tell ☐
9. Do the benefits of the experimental intervention outweigh the harms and costs? *CONSIDER:* • *What was the size of the intervention or treatment effect?* • *Were harms or unintended effects reported for each study group?* • *Was a cost-effectiveness analysis undertaken? (Cost-effectiveness analysis allows a comparison to be made between different interventions used in the care of the same condition or problem.)*	Yes ☐	No ☐	Can't tell ☐

Section D: Will the results help locally?			
10. **Can the results be applied to your local population/in your context?** *CONSIDER:* • *Are the study participants similar to the people in your care?* • *Would any differences between your population and the study participants alter the outcomes reported in the study?* • *Are the outcomes important to your population?* • *Are there any outcomes you would have wanted information on that have not been studied or reported?* • *Are there any limitations of the study that would affect your decision?*	Yes ☐	No ☐	Can't tell ☐
11. **Would the experimental intervention provide greater value to the people in your care than any of the existing interventions?** *CONSIDER:* • *What resources are needed to introduce this intervention taking into account time, finances, and skills development or training needs?* • *Are you able to disinvest resources in one or more existing interventions in order to be able to re-invest in the new intervention?*	Yes ☐	No ☐	Can't tell ☐

APPRAISAL SUMMARY: *Record key points from your critical appraisal in this box. What is your conclusion about the paper? Would you use it to change your practice or to recommend changes to care/interventions used by your organisation? Could you judiciously implement this intervention without delay?*

Appendix E

Reasons for Exclusion of Studies

Author(s)	Year	Title	Exclusion Criteria
Aagaard et al.	2017	The efficacy of psychoeducation on recurrent depression: a randomized trial with a 2-year follow-up	no predictor/moderator analysis as per inclusion criteria
Abbass	2006	Intensive Short-Term Dynamic Psychotherapy of treatment-resistant depression: a pilot study	no predictor/moderator analysis as per inclusion criteria
Abel et al.	2016	Sudden Gains in Treatment Resistant Depression	no predictor/moderator analysis as per inclusion criteria
Abel et al.	2013	Cognitive-behavioral therapy improved response and remission at 6 and 12 months in treatment-resistant depression	wrong article type (commentary)
Andrews et al.	2020	Sudden Gains and Patterns of Symptom Change in Cognitive–Behavioral Therapy for Treatment-Resistant Depression	no predictor/moderator analysis as per inclusion criteria
Arnow et al.	2007	Dropouts versus completers among chronically depressed outpatients	no predictor/moderator analysis as per inclusion criteria
Aust et al.	2022	Efficacy of Augmentation of Cognitive Behavioral Therapy With Transcranial Direct Current Stimulation for Depression	participant criteria as per inclusion criteria not met
Barnhofer et al.	2009	Mindfulness-based cognitive therapy as a treatment for chronic depression: A preliminary study	participant criteria as per inclusion criteria not met
Bausch et al.	2017	Cognitive Behavioral Analysis System of Psychotherapy versus Escitalopram in Patients with Chronic Depression: Results from a Naturalistic Long-Term Follow-Up	wrong article type (letter to editor)
Bausch et al.	2017	The impact of childhood maltreatment on the differential efficacy of CBASP versus escitalopram in patients with chronic depression: A secondary analysis	participant criteria as per inclusion criteria not met
Beddig et al.	2020	Mindfulness-based focused attention training versus progressive muscle relaxation in remitted depressed patients: Effects on salivary cortisol and associations with subjective improvements in daily life	participant criteria as per inclusion criteria not met
Beutel et al.	2022	Recovery from chronic depression and structural change: 5-year outcomes after psychoanalytic and cognitive-behavioural long-term treatments (lac depression study)	participant criteria as per inclusion criteria not met
Blackburn & Moore	1997	Controlled acute and follow-up trial of cognitive therapy and pharmacotherapy in out-patients with recurrent depression	no predictor/moderator analysis as per inclusion criteria
Blalock et al.	2008	Cognitive and behavioral mediators of combined pharmacotherapy and psychotherapy of chronic depression.	no predictor/moderator analysis as per inclusion criteria
Bollman et al.	2015	Psychotherapy in old age: The Cognitive Behavioral Analysis System of Psychotherapy (CBASP) for chronically depressed elderly patients	inpatient setting

Bowie et al.	2013	Cognitive remediation for treatment-resistant depression: Effects on cognition and functioning and the role of online homework	no predictor/moderator analysis as per inclusion criteria
Chiesa et al.	2015	Mindfulness-based cognitive therapy vs. psychoeducation for patients with major depression who did not achieve remission following antidepressant treatment	no predictor/moderator analysis as per inclusion criteria
Cladder-Micus et al.	2019	Effects of mindfulness-based cognitive therapy on a behavioural measure of rumination in patients with chronic, treatment-resistant depression	no predictor/moderator analysis as per inclusion criteria
Conradi et al.	2008	Cognitive–behavioural therapy v. usual care in recurrent depression	participant criteria as per inclusion criteria not met
Constantino et al.	2012	The relation between changes in patients' interpersonal impact messages and outcome in treatment for chronic depression	no predictor/moderator analysis as per inclusion criteria
Constantino et al.	2016	Change in Patients' Interpersonal Impacts as a Mediator of the Alliance-Outcome Association in Treatment for Chronic Depression	no predictor/moderator analysis as per inclusion criteria
Corney et al.	2005	Thirty-six-month outcome data from a trial of counselling with chronically depressed patients in a general practice setting	participant criteria as per inclusion criteria not met
DeMello et al.	2001	A randomized controlled trial comparing moclobemide and moclobemide plus interpersonal psychotherapy in the treatment of dysthymic disorder	participant criteria as per inclusion criteria not met
Den Boer et al.	2007	Cognitive self-therapy for chronic depression and anxiety: A multi-centre randomized controlled study	participant criteria as per inclusion criteria not met
D'Urso et al.	2013	Transcranial Direct Current Stimulation and Cognitive-Behavioral Therapy: Evidence of a Synergistic Effect in Treatment-Resistant Depression	wrong article type (letter to editor/ case report)
Eisendrath et al.	2008	Mindfulness-based cognitive therapy for treatment-resistant depression: A pilot study	no predictor/moderator analysis as per inclusion criteria
Feldman et al.	2014	Mindfulness based cognitive therapy versus a health enhancement program for treatment resistant depression: a randomized controlled trial	wrong article type (study protocol)
Fonagy et al.	2015	Pragmatic randomized controlled trial of long-term psychoanalytic psychotherapy for treatment-resistant depression: The Tavistock Adult Depression Study (TADS)	no predictor/moderator analysis as per inclusion criteria
Forkmann et al.	2016	The Effects of Mindfulness-Based Cognitive Therapy and Cognitive Behavioral Analysis System of Psychotherapy added to Treatment as Usual on suicidal ideation in chronic depression: Results of a randomized-clinical trial	no predictor/moderator analysis as per inclusion criteria
Foroughi et al.	2020	The effectiveness of mindfulness-based cognitive therapy for reducing rumination and improving mindfulness and self-compassion in patients with treatment-resistant depression	no predictor/moderator analysis as per inclusion criteria
Frank et al.	2007	Randomized trial of weekly, twice-monthly, and monthly interpersonal psychotherapy as maintenance treatment for women with recurrent depression	no predictor/moderator analysis as per inclusion criteria

Friedman et al.	2009	Cognitive therapy augmentation versus CT switch treatment: A STAR*D report	no predictor/moderator analysis as per inclusion criteria
Furukawa et al.	2018	Cognitive-Behavioural Analysis System of Psychotherapy (CBASP), a drug, or their combination: differential therapeutics for persistent depressive disorder: a study protocol of an individual participant data network meta-analysis	participant criteria as per inclusion criteria not met
Graser et al.	2006	Effects of a 12-Week Mindfulness, Compassion, and Loving Kindness Program on Chronic Depression: A Pilot Within-Subjects Wait-List Controlled Trial	no predictor/moderator analysis as per inclusion criteria
Harley et al.	2008	Adaptation of dialectical behavior therapy skills training group for treatment-resistant depression	no predictor/moderator analysis as per inclusion criteria
Humer et al.	2020	Effects of alliance ruptures and repairs on outcomes	no predictor/moderator analysis as per inclusion criteria
Jarrett et al.	2001	Preventing recurrent depression using cognitive therapy with and without a continuation phase: A randomized clinical trial	participant criteria as per inclusion criteria not met
Keller et al.	2000	A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression	no predictor/moderator analysis as per inclusion criteria
Klein et al.	2003	Therapeutic Alliance in Depression Treatment: Controlling for Prior Change and Patient Characteristics.	no predictor/moderator analysis as per inclusion criteria
Klein et al.	2011	Psychotherapy of chronic depression with cognitive behavioral analysis system of psychotherapy (CBASP)	wrong article type (overview of intervention)
Kocsis et al.	2009	Cognitive behavioral analysis system of psychotherapy and brief supportive psychotherapy for augmentation of antidepressant nonresponse in chronic depression: The REVAMP trial	no predictor/moderator analysis as per inclusion criteria
Kuyken et al.	2008	Mindfulness-Based Cognitive Therapy to Prevent Relapse in Recurrent Depression	participant criteria as per inclusion criteria not met
Kuyken et al.	2015	The effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse/recurrence: Results of a randomised controlled trial (The PREVENT study)	participant criteria as per inclusion criteria not met
Lau	2020	Mindfulness-based cognitive therapy: A low intensity group program to prevent depressive relapse	wrong article type (book chapter)
Ledari et al.	2018	A Comparison Between the Effectiveness of Acceptance and Commitment Treatment and Behavioral Activation Treatment for Depression on Symptoms Severity and Rumination Among Patients with Treatment-Resistant Depression	no predictor/moderator analysis as per inclusion criteria
Leuzinger-Bohleber	2017	'Consenting to be robbed so as not to be murdered': psychoanalytic treatments of chronically depressed patients in two parallel depression research studies	wrong article type (conference abstract)

Leuzinger-Bohleber et al.	2019	Outcome of psychoanalytic and cognitive-behavioural long-term therapy with chronically depressed patients: A controlled trial with preferential and randomized allocation	participant criteria as per inclusion criteria not met
Leuzinger-Bohleber et al.	2020	The LAC Study: A comparative outcome study of psychoanalytic and cognitive-behavioral long-term therapies of chronic depressive patients	wrong article type (book chapter)
Lo et al.	2015	Evaluating compassion–mindfulness therapy for recurrent anxiety and depression: A randomized control trial.	participant criteria as per inclusion criteria not met
Matsunaga et al.	2010	Psychosocial functioning in patients with treatment-resistant depression after group cognitive behavioral therapy	no predictor/moderator analysis as per inclusion criteria
McCullough et al.	1997	Cognitive-Behavior Therapy for Chronic Depression (CBT-CD): Combined national collaborative study	wrong article type (conference abstract)
McCullough	2003	Treatment for chronic depression using Cognitive Behavioral Analysis System of Psychotherapy (CBASP)	wrong article type
McLoughlin et al.	2021	Mindfulness based cognitive therapy for recurrent depressive disorder	wrong article type
Meister et al.	2020	Adverse events during a disorder-specific psychotherapy compared to a nonspecific psychotherapy in patients with chronic depression	participant criteria as per inclusion criteria not met
Melyani et al.	2015	Mindfulness based cognitive therapy versus cognitive behavioral therapy in cognitive reactivity and self-compassion in females with recurrent depression with residual symptoms	not in English or German language
Michalak et al.	2016	Mindfulness-Based Cognitive Therapy and a Group Version of the Cognitive Behavioral Analysis System of Psychotherapy for Chronic Depression: Follow-Up Data of a Randomized Controlled Trial and the Moderating Role of Childhood Adversities	wrong article type (letter to editor)
Michalak et al.	2015	A randomized controlled trial on the efficacy of mindfulness-based cognitive therapy and a group version of cognitive behavioral analysis system of psychotherapy for chronically depressed patients	no predictor/moderator analysis as per inclusion criteria
Minelli et al.	2019	Clinical efficacy of trauma-focused psychotherapies in treatment-resistant depression (TRD) in-patients: A randomized, controlled pilot-study	no predictor/moderator analysis as per inclusion criteria
Moeller et al.	2020	Rumination-focused cognitive behaviour therapy for non-responsive chronic depression: An uncontrolled group study	no predictor/moderator analysis as per inclusion criteria
Morriss et al.	2016	Efficacy and cost-effectiveness of a specialist depression service versus usual specialist mental health care to manage persistent depression: a randomised controlled trial	no predictor/moderator analysis as per inclusion criteria
Monnart et al.	2019	Treatment of Resistant Depression: A Pilot Study Assessing the Efficacy of a tDCS-Mindfulness Program Compared With a tDCS-Relaxation Program	no predictor/moderator analysis as per inclusion criteria
Morimoto et al.	2014	Neuroplasticity-based computerized cognitive remediation for treatment-resistant geriatric depression	no predictor/moderator analysis as per inclusion criteria

Murray et al.	2010	Relief of Chronic or Resistant Depression (Re-ChORD): A pragmatic, randomized, open-treatment trial of an integrative program intervention for chronic depression	no predictor/moderator analysis as per inclusion criteria
Nakagawa et al.	2017	Effectiveness of Supplementary Cognitive-Behavioral Therapy for Pharmacotherapy-Resistant Depression: A Randomized Controlled Trial	no predictor/moderator analysis as per inclusion criteria
Ninan et al.	2002	Symptomatic and syndromal anxiety in chronic forms of major depression: Effect of nefazodone, Cognitive Behavioral Analysis System of Psychotherapy, and their combination	no predictor/moderator analysis as per inclusion criteria
O'Mahen et al.	2019	Trajectories of Change in a Group Behavioral Activation Treatment for Severe, Recurrent Depression	participant criteria as per inclusion criteria not met
Ostacoli et al.	2018	Comparison of eye movement desensitization reprocessing and cognitive behavioral therapy as adjunctive treatments for recurrent depression: The European Depression EMDR Network (EDEN) randomized controlled trial	no predictor/moderator analysis as per inclusion criteria
Papageorgiou et al.	2015	Group Metacognitive Therapy for Severe Antidepressant and CBT Resistant Depression: A Baseline-Controlled Trial	no predictor/moderator analysis as per inclusion criteria
Renner et al.	2018	Exploring mechanisms of change in schema therapy for chronic depression	no predictor/moderator analysis as per inclusion criteria
Roehricht et al.	2013	An exploratory randomized controlled trial of body psychotherapy for patients with chronic depression	no predictor/moderator analysis as per inclusion criteria
Sayegh et al.	2012	Cognitive behavioural analysis system of psychotherapy for treatment-resistant depression: Adaptation to a group modality	no predictor/moderator analysis as per inclusion criteria
Schanche et al.	2021	Self-criticism and self-reassurance in individuals with recurrent depression: Effects of Mindfulness-Based Cognitive Therapy and relationship to relapse	participant criteria as per inclusion criteria not met
Schnitzler & Christenhusz	2016	Effectiveness of MBCT in addition to treatment as usual in a patient group with chronic anxiety or depression: A pilot-study	not in English or German language
Schramm et al.	2015	From animal behavior to human health an animal-assisted mindfulness intervention for recurrent depression	no predictor/moderator analysis as per inclusion criteria
Schramm et al.	2017	Effect of disorder-specific vs nonspecific psychotherapy for chronic depression: A randomized clinical trial	no predictor/moderator analysis as per inclusion criteria
Schuling et al.	2018	The Co-creation and Feasibility of a Compassion Training as a Follow-up to Mindfulness-Based Cognitive Therapy in Patients with Recurrent Depression	no predictor/moderator analysis as per inclusion criteria
Schuling et al.	2020	Recovery from recurrent depression: Randomized controlled trial of the efficacy of mindfulness-based compassionate living compared with treatment-as-usual on depressive symptoms and its consolidation at longer term follow-up	participant criteria as per inclusion criteria not met

Schearer et al.	2019	Refractory depression - cost-effectiveness of radically open dialectical behaviour therapy: findings of economic evaluation of RefraMED trial	no predictor/moderator analysis as per inclusion criteria
Scott	1992	Chronic Depression: Can Cognitive Therapy Succeed When Other Treatments Fail?	no predictor/moderator analysis as per inclusion criteria
Simpson et al.	2003	A randomized controlled trial to evaluate the effectiveness and cost-effectiveness of psychodynamic counselling for general practice patients with chronic depression	participant criteria as per inclusion criteria not met
Simpson et al.	2000	A randomised controlled trial to evaluate the effectiveness and cost-effectiveness of counselling patients with chronic depression	participant criteria as per inclusion criteria not met
Sledge	1999	Treatment of chronic depression with ICBT-CD: A preliminary treatment-outcome study assessing interpersonal and attributional correlates of chronic depression	could not access
Souza et al.	2016	Interpersonal psychotherapy as add-on for treatment-resistant depression: A pragmatic randomized controlled trial	participant criteria as per inclusion criteria not met
Stalsett et al.	2012	Existential dynamic therapy ("VITA") for treatment-resistant depression with Cluster C disorder: Matched comparison to treatment as usual	inpatient setting
Strauss et al.	2012	Group person-based cognitive therapy for chronic depression: A pilot randomized controlled trial	no predictor/moderator analysis as per inclusion criteria
Swan et al.	2004	Coping with depression: An open study of the efficacy of a group psychoeducational intervention in chronic, treatment-refractory depression	no predictor/moderator analysis as per inclusion criteria
Swan et al.	2014	Cognitive Behavioural Analysis System of Psychotherapy (CBASP) for chronic depression: Clinical characteristics and six month clinical outcomes in an open case series	no predictor/moderator analysis as per inclusion criteria
Ter Avest et al.	2021	Prospective associations between home practice and depressive symptoms in mindfulness-based cognitive therapy for recurrent depression: A 15 months follow-up study	participant criteria as per inclusion criteria not met
Ter Avest et al.	2021	Interplay between self-compassion and affect during Mindfulness-Based Compassionate Living for recurrent depression: An Autoregressive Latent Trajectory analysis	participant criteria as per inclusion criteria not met
Thase et al.	1994	Response to cognitive-behavioral therapy in chronic depression	could not access
Town et al.	2017	A randomised controlled trial of Intensive Short-Term Dynamic Psychotherapy for treatment resistant depression: the Halifax Depression Study	no predictor/moderator analysis as per inclusion criteria
Uebelacker et al.	2012	Adapted Behavior Therapy for Persistently Depressed Primary Care Patients An Open Trial	no predictor/moderator analysis as per inclusion criteria
Van Aalderen et al.	2012	The efficacy of mindfulness-based cognitive therapy in recurrent depressed patients with and without a current depressive episode: A randomized controlled trial	no predictor/moderator analysis as per inclusion criteria

Valenstein et al.	2016	Augmenting Ongoing Depression Care With a Mutual Peer Support Intervention Versus Self-Help Materials Alone: A Randomized Trial	Not psychological intervention
Van Aalderen et al.	2015	Long-term outcome of mindfulness-based cognitive therapy in recurrently depressed patients with and without a depressive episode at baseline	no predictor/moderator analysis as per inclusion criteria
Vittengl et al.	2004	Self-directed affiliation and autonomy across acute and continuation phase cognitive therapy for recurrent depression	no predictor/moderator analysis as per inclusion criteria
Vittengl et al.	2010	Improvement in social-interpersonal functioning after cognitive therapy for recurrent depression	no predictor/moderator analysis as per inclusion criteria
Vittengl et al.	2009	Deterioration in psychosocial functioning predicts relapse/recurrence after cognitive therapy for depression	no predictor/moderator analysis as per inclusion criteria
Vittengl et al.	2019	Estimating outcome probabilities from early symptom changes in cognitive therapy for recurrent depression	no predictor/moderator analysis as per inclusion criteria
Vittengl et al.	2022	Stability and Change in Relations Between Personality Traits and the Interpersonal Problems Circumplex During Cognitive Therapy for Recurrent Depression	no predictor/moderator analysis as per inclusion criteria
Vittengl et al.	2022	Does Symptom Linkage Density Predict Outcomes in Cognitive Therapy for Recurrent Depression?	no predictor/moderator analysis as per inclusion criteria
Watkins et al.	2011	Rumination-focused cognitive-behavioural therapy for residual depression: Phase II randomised controlled trial	participant criteria as per inclusion criteria not met
Watkins et al.	2011	An effectiveness trial of group cognitive behavioral therapy for patients with persistent depressive symptoms in substance abuse treatment	participant criteria as per inclusion criteria not met
Weck et al.	2013	The Relationship Between Therapist Competence and Homework Compliance in Maintenance Cognitive Therapy for Recurrent Depression: Secondary Analysis of a Randomized Trial	participant criteria as per inclusion criteria not met
Wells et al.	2012	Metacognitive therapy in treatment-resistant depression: A platform trial	no predictor/moderator analysis as per inclusion criteria
Werner et al.	2018	A cluster randomized controlled platform trial comparing group MEmory specificity training (MEST) to group psychoeducation and supportive counselling (PSC) in the treatment of recurrent depression	participant criteria as per inclusion criteria not met
Wiersma et al.	2014	The effectiveness of the cognitive behavioral analysis system of psychotherapy for chronic depression: A randomized controlled trial	no predictor/moderator analysis as per inclusion criteria
Wiles et al.	2013	Cognitive behavioural therapy as an adjunct to pharmacotherapy for primary care based patients with treatment resistant depression: Results of the CoBalT randomised controlled trial	no predictor/moderator analysis as per inclusion criteria
Wiles et al.	2016	Long-term effectiveness and cost-effectiveness of cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care: Follow-up of the CoBalT randomised controlled trial	no predictor/moderator analysis as per inclusion criteria

Winnebeck et al.	2017	Brief training in mindfulness meditation reduces symptoms in patients with a chronic or recurrent lifetime history of depression: A randomized controlled study	no predictor/moderator analysis as per inclusion criteria
Wong	2008	Cognitive behavioral treatment groups for people with chronic depression in Hong Kong: A randomized wait-list control design	participant criteria as per inclusion criteria not met
Wong	2009	A six-month follow-up study of cognitive-behavioural treatment groups for Chinese people with depression in Hong Kong	participant criteria as per inclusion criteria not met
Yang et al.	2005	Effect of psychosocial intervention on quality of life and other factors in patients with recurrent depression	not in English or German language
Yasinski et al.	2020	Processes of change in cognitive behavioral therapy for treatment-resistant depression: Psychological flexibility, rumination, avoidance, and emotional processing	no predictor/moderator analysis as per inclusion criteria
Yeon-Hee & Byun	2017	Therapeutic Mechanism of MBCT and Clinical Application of MBCT(Mindfulness-Based Cognitive Therapy) Program on Chronic Depression	not in English or German language

Appendix F

Overview of Intervention Characteristics

Table F1

Overview of Intervention Characteristics

Study	Intervention	Comparison Condition(s)	Intervention Session *N* / Duration	Follow Up	ADM allowed	Intervention Efficacy
Cladder-Micus et al. (2018)	MBCT (group)	TAU	8/8W	3M, 6M	Yes	small to medium effect size (d = 0.35) for MBCT completers only
Eisendrath et al. (2016)	MBCT (group)	HEP + TAU	8/8W	24W, 36W, 52W	Yes	MBCT significantly % higher reduction in depression symptoms, no significant difference in remission rates
Lopez & Basco (2015)	CBT	TAU	18/18W	No	Yes	CBT significant higher remission and clinically significant response rate
Potijk et al. (2020)	CBASP	None	≥30/48W	No	Yes	medium effect size for CBASP at 12M (d = 0.51)
Renner & Berry (2011)	CBT (group)	Self-Help Group / WL	NR/16W	No	NR	Significant symptoms change on one of two outcome measures for CBT
Stangier et al. (2021)	MBT (group) + CBT (individual)	WL	17/NR	1M, 6M	Yes	High effect size for MBT+CBT
Taubner et al. (2011)	LTPP	Healthy Controls	NR/60W	No	No	High effect size for LTPP
Secondary Data Analysis to Fonagy et al. (2015)						
Rost et al. (2019)	LTPP	TAU	60/60W	24M, 30M, 48M	Yes	LTPP greater symptom reduction, p = .017
Secondary Data Analysis to Keller et al. (2000)						
Arnow et al. (2003) **Denton et al. (2010)** **Kocsis et al. (2009a)**	CBASP	ADM / ADM + CBASP	16/12W	No	ADM conditions only	Combined treatment (CBASP+ADM) had significantly higher response rate

Manber et al. (2008) **Nemeroff et al. (2003)**	CBASP	ADM / ADM + CBASP	16/12W	No	ADM conditions only	Combined treatment (CBASP+ADM) had significantly higher response rate
Secondary Data Analysis to Phase 2 of Koscis et al. (2009b)						
Arnow et al. (2013) **Schankman et al. (2013)** **Steidtmann et al. (2012)**	CBASP	BSP BSP / ADM BSP / ADM	16-20/12W	No	Yes	Similar response rates across all three conditions, Significant symptom reduction across all conditions
Secondary Data Analysis to Michalak et al. (2015)						
Probst et al. (2020)	MBCT (group)	CBASP (group)/ TAU	NR/8W	6M	Yes	No significant difference in remission rates between MBCT and CBASP, CBASP and MBCT more effective than TAU
Secondary Data Analysis to Schramm et al. (2017) and Schramm et al. (2019)						
Assmann et al. (2018) **Bausch et al. (2020)** **Erkens et al. (2018)** **Klein et al. (2018)** **Serbanescu et al. (2020)**	CBASP	SP	32/48W	No 12M, 24M No No No	No	CBASP significantly more effective than SP with small to medium effect size (d = 0.31)
Secondary Data Analysis to Wiles et al. (2013)						
Button et al. (2015)	CBT	TAU	≤ 18/NR	No	Yes	Significantly higher response rate for CBT

Note. Antidepressant medication (ADM), Brief Supportive Psychotherapy (BSP), Cognitive Behavioural Analysis System of Psychotherapy (CBASP), Health Enhancement Programme (HEP), Long Term Psychoanalytic Psychotherapy (LTPP), Months (M), Mindfulness-based Cognitive Therapy (MBCT), Metta-Based Therapy (MBT), Not reported (NT), Supportive Psychotherapy (SP), Treatment as usual (TAU), Weeks (W), Wait List (WL)

Appendix G

Summary of Risk of Bias Assessments

Table G1

Overview Risk of Bias Assessment – Randomised Control Trials

Study	*Focused research question?*	*Appropriate Randomised?*	*Are all participants accounted for?*	*Participants and investigators blinded?*	*Are study groups similar?*	*Similar care across groups (except for experimental intervention)?*	*Effects of intervention reported comprehensively?*	*Precision of treatment effect reported?*	*Benefits greater than harm/ cost?*	*Results applicable to local population/ your context?*	*Experimental intervention greater value than existing interventions?*	***Risk of Bias***
Cladder-Micus et al. (2018)	Y	Y	Y	N	Y	Y	P	Y	Y	Y	N	**MEDIUM**
Eisendrath et al (2016)	Y	Y	Y	N	Y	Y	P	N	CT	Y	CT	**MEDIUM**
Renner & Berry (2011)	Y	Y	Y	N	CT	CT	P	N	N	N	N	**HIGH**
Stangier et al. (2021)	Y	Y	Y	P	P	Y	Y	Y	Y	Y	CT	**MEDIUM**
Secondary data analysis to Fonagy et al. (2015)												
Rost et al. (2019)	Y	Y	Y	P	P	Y	Y	Y	Y	Y	CT	**MEDIUM**
Secondary data analysis to Keller et al. (2000)												
Arnow et al. (2003)	Y	Y	Y	P	Y	Y	P	Y	Y	Y	Y	**LOW**

Denton et al. (2010)	Y	Y	Y	P	Y	Y	P	Y	Y	Y	Y	**LOW**
Kocsis et al. (2009a)	Y	Y	Y	P	Y	Y	P	Y	Y	Y	Y	**LOW**
Manber et al. (2008)	Y	Y	Y	P	Y	Y	P	Y	Y	Y	Y	**LOW**
Nemeroff et al. (2003)	Y	Y	Y	P	Y	Y	P	Y	Y	Y	Y	**LOW**
Secondary Data Analysis to Kocsis et al. (2009b)												
Arnow et al. (2013)	Y	Y	Y	P	Y	Y	P	Y	Y	Y	Y	**LOW**
Schankman et al. (2013)	Y	Y	Y	P	Y	Y	P	Y	Y	Y	Y	**LOW**
Steidtman et al. (2012)	Y	Y	Y	P	Y	Y	P	Y	Y	Y	Y	**LOW**
Secondary data analysis to Michalak et al. (2015)												
Probst et al. (2020)	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	CT	**LOW**
Secondary data analysis to Schramm et al. (2017) and Schramm et al. (2019)												
Assmann et al. (2018)	Y	Y	Y	P	P	Y	Y	Y	CT	Y	Y	**MEDIUM**
Bausch et al. (2020)	Y	Y	Y	P	P	Y	Y	Y	CT	Y	Y	**MEDIUM**
Erkens et al. (2018)	Y	Y	Y	P	P	Y	Y	Y	CT	Y	Y	**MEDIUM**
Klein et al. (2018)	Y	Y	Y	P	P	Y	Y	Y	CT	Y	Y	**MEDIUM**
Serbanescu et al. (2020)	Y	Y	Y	P	P	Y	Y	Y	CT	Y	Y	**MEDIUM**

Secondary data analysis to Wiles et al. (2013)												
Button et al. (2015)	Y	Y	Y	P	CT	Y	Y	Y	Y	Y	N	**MEDIUM**

Note. N – No (item not adequately addressed), Y – Yes (item adequately addressed), CT – Can't tell if item adequately addressed, P – Partially (Item partially addressed).

Table G2

Overview Risk of Bias Assessment – Case-Control Studies

Study	*Clearly focused research question?*	*Appropriate study method?*	*Cases recruited in an acceptable way?*	*Controls recruited in an acceptable way?*	*Was the exposure accurately measured to minimise bias?*	*Aside from experimental intervention, were groups treated equally?*	*Were confounding factors considered in design/ analysis?*	*Treatment effect size?*	*Precision estimate of treatment effect reported?*	*Do you believe the results?*	*Results applicable to local population/ your context?*	*Do results fit with other available evidence?*	**Risk of Bias**
Lopez & Basco (2015)	Y	Y	Y	Y	P	Y	Y	Medium to large effect size	Y	Y	Y	Y	**LOW**
Taubner et al. (2011)	Y	Y	Y	Y	P	N	CT	High	P	CT	N	CT	**HIGH**

Note. N – No (item not adequately addressed), Y – Yes (item adequately addressed), CT – Can't tell if item adequately addressed, P – Partially (Item partially addressed).

Table G3

Overview Risk of Bias Assessment – Cohort Studies

Study	*Clearly focused research question?*	*Cohort recruited in an acceptable way?*	*Exposure accurately measured to minimise bias?*	*Outcome accurately measured to minimise bias?*	*All important confounding factors identified?*	*Were confounding factors considered in design/ analysis?*	*FU of subjects complete enough?*	*FU of subjects long enough?*	*Study results*	*Were results precise?*	*Do you believe the results?*	*Results applicable to local population/ your context?*	*Do results fit with other available evidence?*	*What are the implications of this study for practice?*	**Risk of Bias**
Potijk et al. (2020)	Y	Y	P	Y	Y	Y	Y	P	CBASP (group) more effective for late-onset than early-onset depression.	Y	Y	Y	CT	Limited due to lack of control group and sample size	**MEDIUM**

Note. N – No (item not adequately addressed), Y – Yes (item adequately addressed), CT – Can't tell if item adequately addressed, P – Partially (Item partially addressed).

Part Two: Empirical Study

Treatment of Complex Cases of Depression in Primary Mental Health Services: A Quantitative Study.

Abstract

Background: Patient responses to psychological treatments for depression are varied. Previous research highlights how the presence and interaction of multiple patient factors can negatively affect treatment outcomes, contributing to increased case complexity. This study examined whether complex cases showed differential treatment responses to two common treatments for depression, namely cognitive behavioural therapy (CBT) and counselling for depression (CfD).

Method: A secondary data analysis of a randomised controlled trial conducted within National Health Service Talking Therapies services was completed. Machine learning algorithms were used in the original trial to classify patients as complex or standard cases based on presence of certain psychosocial characteristics. Only patients (n = 323) who accessed high-intensity CBT of CfD treatment were included in this study. Multiple regression analysis examined the relationship between case complexity and treatment response as measured by the Patient Health Questionnaire-9. Additional sensitivity analyses were completed. Differences in deterioration and remission rates were examined using a chi-square test.

Results: Treatment effect sizes were large for complex cases (Cohen's d = 1.41). No significant main effect of case complexity, treatment modality and the interaction between case complexity and treatment modality on treatment outcomes was found. Complex cases had comparable reliable remission and deterioration rates for CBT and CfD, with over half (55.8%) showing reliable improvement following either intervention.

Conclusion: Complex cases appear to respond similarly to CBT and CfD, with both interventions equally effective in addressing depression symptoms.

Keywords*: Depression, Adults, CBT, Counselling, Treatment Outcomes*

Practitioner Points

- High-intensity CBT and CfD are acceptable treatments for complex cases with depression symptoms, as indicated by large treatment effect sizes.
- A substantial proportion of complex cases do not achieve reliable improvements with high-intensity CBT and CfD. Clinicians should therefore also consider alternative treatment approaches.
- Future research is needed to help understand how depression outcomes for complex cases can be improved further.

Introduction

Depression is one of the most common mental health difficulties in England, with nearly one fifth of adults reporting being diagnosed at some point in their lives (Bridges, 2016). For depression alone, the cost of health services and lost earnings was estimated at £7.5 billion in 2007 and is expected to rise approximately 60% by 2026 (McCrone et al., 2008). Although psychological therapies can contribute to the remission of depression, a meta-analysis showed that only 62% of patients improved following psychotherapy (Cuijpers et al., 2014). Additionally, some patients deteriorate following psychotherapy, and this has also been specifically reported for depression (Cuijpers et al., 2018; Hansen et al., 2002). Cuijpers et al. (2018) conducted a meta-analysis of psychotherapy studies on adult depression and reported a median deterioration rate of 4%, going up to 10% in some studies. These statistics highlight the need to further improve treatment outcomes for depression. This evidence also indicates that there is considerable variability in treatment response, such that some patients respond well to psychotherapy for depression, others do not benefit much from it, while others seem to deteriorate (Cuijpers et al., 2018).

Given the evidence of treatment response variability, research has focused on understanding which factors and patient characteristics may contribute to differential treatment responses in depression. For instance, the experience of childhood maltreatment is linked to poorer outcomes (Nanni et al., 2012). Sociodemographic characteristics such as socioeconomic deprivation, employment and marital status have also been found to be associated with depression treatment outcomes (Buckman et al., 2021a, 2022; Finegan et al., 2018). Patients with comorbid anxiety and chronic health problems also seem to have poorer depression treatment outcomes (Buckman et al., 2021b; Delgadillo et al., 2017). However, the impact of other characteristics, such as personality pathology is less clear due to inconsistent findings across the literature (Banyard et al., 2021; Mulder, 2002; Newton-Howes et al., 2014). Newton-

Howes et al. (2014) concluded in their meta-analysis that adults with comorbid personality disorders have double the odds of an unfavourable outcome than those without such a comorbidity, regardless of offered intervention type. This contrasts with a recent meta-analysis by Banyard et al. (2021), who found that comorbid personality disorder only had a small, negative impact on cognitive behaviour therapy (CBT) outcomes for adults with depression. Further analysis showed that this effect was non-significant when adjusting for baseline depression severity. Banyard et al. (2021) also noted that those with a comorbid personality disorder appear to benefit from CBT treatment with longer duration, highlighting how treatment requirements vary amongst patient groups.

More recently, attempts have been made to consider the impact and interaction of multiple patient characteristics on treatment outcomes in depression, rather than viewing such factors in isolation from each other (see review by Cohen & DeRubeis, 2018). For instance, treatment non-responders may present with characteristics that negatively affect outcomes. Such patient characteristics may contribute to increased case complexity; that is, cases presenting with a range of factors that complicate the course and outcome of treatment.

Previous authors have attempted to conceptualise case complexity, in order to shed light on the cases for whom psychotherapy does not work as well as expected, and for whom some adaptations or targeted interventions may be required. For instance, Ruscio and Holohan (2006) summarised factors often observed in complex cases, summarising them under clinical, psychosocial, motivational, or physical domains. Specific examples of such factors are: severe symptomatology, presence of comorbidities, chronic pain, low educational level, low self-esteem, social isolation, and lack of treatment compliance. Furthermore, healthcare research has attempted to identify complex cases, such as by applying a biopsychosocial model (de Jonge et al., 2005). De Jonge et al. (2005) proposed ‘INTERMED’, a psychometric tool that can be applied flexibly across healthcare to determine case complexity. Twenty variables

covering biological (chronicity, diagnostic dilemma, symptom severity, diagnostic challenge, complications and life threat), psychological (restrictions in coping, psychiatric dysfunction, resistance to treatment, psychiatric symptoms, mental health threat), social (restrictions in integration, social dysfunction, residential instability, restrictions of network, social vulnerability) and health care factors (treatment intensity, treatment experience, organisation of care, appropriateness of referral, coordination) feed into a final case complexity score. The 'INTERMED' tool is found to be of good validity and reliability but has been mainly applied in physical healthcare settings as opposed to mental healthcare settings (Oliveira et al., 2022).

More recently, the cumulative complexity model was proposed by Shippee et al. (2012). Based on the existing evidence base, it proposes that imbalances between patient workload (e.g., daily demands and challenges) and capacity (e.g., physical and emotional health, finances, social support to meet daily challenges) negatively affect patient outcomes and influence the development of case complexity. A key contribution of this theoretical model is the notion that complexity arises from the cumulative (e.g., additive) influence of multiple biopsychosocial characteristics that are statistically associated with their health status and/or treatment response. Within the cumulative complexity model patient workload and patient capacity are proposed to influence each other, as well as patient outcomes directly and indirectly via level of access to care, use of care and self-care. Additionally, patient outcomes are suggested to influence patient workload via 'burden of treatment'. For instance, patients who are showing poorer outcomes may be offered further treatments to address this, thereby increasing patient workload. Patient outcomes are also suggested to influence patient capacity via 'burden of illness'. For example, if there is a further deterioration in health then patient capacity will reduce further as well.

The cumulative complexity model has been applied in psychotherapy research to study complex cases, to identify their most salient characteristics, and to understand how they

respond to psychological intervention (Delgadillo et al., 2017). Using a machine learning approach the authors identified patient-factors that were independently associated with poorer treatment outcomes: unemployment, belonging to an ethnic minority, younger age, high levels of comorbid depression and anxiety symptoms, high functional impairment, as well as the presence of avoidant, suspicious, impulsive, or dependent personality traits. Patients with a combination of these characteristics, classified as more complex cases, had poorer treatment outcomes in primary mental health services compared to other cases. Furthermore, patients classed as complex cases tended to have especially poor treatment outcomes when offered low intensity psychological interventions. Additionally, the presence of several of these factors was noted to have a negative, cumulative effect on treatment outcomes, consistent with the cumulative complexity model.

Optimising treatment allocation for complex cases may potentially aid in improving treatment outcomes. The National Institute for Health and Care Excellence (NICE) guidelines recommend several psychotherapies for depression, including CBT and person-centred counselling for moderate or severe cases (NICE, 2022). However, emerging research highlights that patients with certain characteristics may respond better to one treatment approach than the other (see review on this topic by Cohen & DeRubeis, 2018). This was also noted in a study by Lorenzo-Luaces et al. (2017) who found that patients with poorer prognosis (e.g., more complex cases) responded better to CBT than other interventions. Currently, treatment allocation is predominantly informed by clinical guidelines and decided by clinical judgement. However, clinical judgment can be prone to bias, such as the clinician's preference for certain theoretical models, information processing biases or overreliance on clinical interviews (Bell & Mellor, 2009). Increasing evidence also highlights that clinical judgement is not very accurate in predicting whether patients deteriorate during interventions (Hannan et al., 2005), and that other approaches such as statistical prediction are more precise (Grove et al., 2000).

Statistically guided or "data-driven" treatment allocation for individuals with depression has also been found superior to clinical judgement in retrospective analyses of clinical trials and routine care interventions (e.g., see Delgadillo & Gonzalez Salas Duhne, 2020; van Bronswijk et al., 2021). More recently, two randomised controlled trials have shown that data-driven allocation of patients to different types of interventions (Delgadillo et al., 2022) or different therapists (Constantino et al., 2021) improves clinical outcomes. As has been argued by Cohen and DeRubeis (2018), statistical prediction may improve treatment outcomes in depression whilst using currently available therapies more efficiently. The literature into this area is growing, but few studies have focused on the role of case complexity in depression treatment selection. Given that complex cases of depression tend to have poorer outcomes in primary mental health settings and that patients with poorer prognosis may respond better to certain psychotherapies, further research is warranted (Delgadillo et al., 2017; Lorenzo-Luaces et al., 2017). According to a retrospective cohort study that identified subgroups of depressed patients who responded better to either CBT or person-centred counselling, the results suggest that more complex cases may respond better to CBT (Delgadillo & Gonzalez Salas Duhne, 2020). However, such findings have not yet been replicated or validated in prospective studies or clinical trials. Therefore, this study builds on the existing literature by focusing on improving understanding of such differential treatment responses for complex cases.

Aims

The overarching aim of the current study was to understand whether complex cases respond differently to two distinctive and routinely available high-intensity psychological treatments for depression, CBT and counselling for depression (CfD). Although a wide range of other psychological interventions are also recommended for depression (NICE, 2022), CBT and CfD were focused on as they are both commonly offered to patients in England within primary mental health services, particularly within NHS Talking Therapies (National Collaborating

Centre for Mental Health, 2023). Focusing on two interventions that are likely to be offered in the first instance to patients in England for the treatment of depression increases relevance of findings to the local population and is more likely to inform future care within the NHS. The present study applied the data-driven definition of complex cases proposed by Delgadillo et al. (2017), which is informed by the cumulative complexity theoretical model (Shippee et al., 2012). The following research questions guided the study:

- Do complex cases identified *a priori* (e.g., based on their psychosocial characteristics, before they started their therapy) respond differently to CBT and CfD based on post-treatment levels of depression symptoms?
- Do complex cases respond differently to CBT and CfD based on remission rates and deterioration rates?

Although the literature in this area is limited, based on findings of a study by Delgadillo and Gonzalez Salas Duhne (2020), it was hypothesised that:

- Case complexity will have a main effect on treatment outcomes (e.g., complex cases were expected to have poorer treatment outcomes compared to other cases).
- Complex cases who received CBT will have better outcomes than complex cases who received CfD.

Clinical Implications

Investigating differential treatment response amongst complex cases could support clinicians in matching patients to the most appropriate and effective treatment to meet their needs. This could help to reduce the impact of bias when solely using clinical judgement in treatment allocation, and thereby potentially improving treatment outcomes.

Method

The American Psychological Association (APA) Style Journal Article Reporting Standards for quantitative studies was used as a guideline for this empirical study (Appendix A; APA, 2020).

Design

This is a secondary analysis of data from a randomised controlled trial conducted across four National Health Service (NHS) Talking Therapies services, formerly known as Improving Access to Psychological Therapies, in England between 2018-2020. The trial and services are described below. For further information on the trial see Delgadillo et al. (2022).

Service Setting and Interventions

Primary mental health services in England are provided by NHS Talking Therapies. NHS Talking Therapies services offer low-intensity treatments (LIT) and high-intensity treatments (HIT) for individuals with symptoms of depression and other common mental health difficulties. NHS Talking Therapies follow a stepped care model, where patients usually receive LITs in the first instance. If this is unsuccessful, patients are stepped up to HITs (National Collaborating Centre for Mental Health, 2023). Treatment allocation is determined by the clinician completing an initial assessment, following clinical guidelines for stepped care.

LIT for depression includes guided self-help based on CBT, computerised CBT, and behavioural activation. These interventions are facilitated by qualified psychological wellbeing practitioners and can be delivered in different formats and settings (e.g., individual, groups, telephone, face to face; National Collaborating Centre for Mental Health, 2023). NHS Talking Therapies provide a range of different HITs for depression, including CBT and CfD. Interventions offered are evidence-based and are in line with NICE guidance (NICE, 2009; NICE, 2022). HITs are facilitated by accredited practitioners who have completed training for the given intervention as specified by the National Collaborating Centre for Mental Health

(2023). For this study, HIT therapists were trained following the national curriculum for high-intensity CBT interventions or high-intensity CfD (Health Education England, 2022; Hill, 2011). To ensure treatment fidelity and competence, all therapists accessed weekly clinical supervision with a clinician highly experienced in the relevant therapeutic modality.

CBT is a goal-oriented and highly structured intervention which focuses on understanding the interaction of thoughts, feelings, and behaviour in order to implement effective coping strategies (NICE, 2022). Within NHS Talking Therapies a treatment manual is followed as recommended by NICE guidelines (2022). The protocol follows either disorder specific Beckian CBT or Martell's Behavioural Activation (Beck, 1979; Martell et al., 2001). Additionally, the core competencies and activities listed by Roth and Pilling (2008) are adhered to.

CfD is a person-centred and experiential therapy based on humanistic approaches (Sanders & Hill, 2014). More specifically, CfD draws on and combines elements of person-centred therapy and emotion-focused therapy (Elliott, 2004; Mearns et al., 2013). Within NHS Talking Therapies, CfD is delivered following evidence-based and standardised treatment protocols to ensure compliance with national competencies (Hill, 2011; Sanders & Hill, 2014). For both high-intensity CBT and CfD, patients can access up to 20 sessions (60-minute session length), which is in line with NICE guidance (2022).

Data Source

The fully anonymised dataset was obtained from the StratCare randomised control trial (Delgadillo et al., 2022). The StratCare trial was conducted in NHS Talking Therapy services managed by Lancashire and South Cumbria NHS Foundation Trust and Rotherham Doncaster and South Humber NHS Foundation Trust. The trial included a total of 951 patients who accessed LITs and HITs. The aim was to explore whether stratified care, namely allocating

patients to treatment based on baseline characteristics, resulted in better treatment outcomes than the usual stepped care approach. Participating therapists (n = 30) were randomised using a computer-generated schedule to either the stratified or stepped care condition. Randomisation was completed by an independent researcher. Therapists were informed of the allocated condition. Patients who accessed NHS Talking Therapies for treatment of a common mental health problem (e.g., depression or anxiety) were classified as standard or complex cases following an initial assessment session and prior to treatment allocation. A machine learning algorithm was used to calculate level of complexity based on baseline psychosocial characteristics (level of depression, anxiety and functional impairment, personality traits, race, ethnicity, and employment status). Within the stratified care condition standard cases were allocated to LITs and complex cases to HITs. Participating patients were blinded to how treatment allocation was determined. Results showed that patient outcomes, as measured by the Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001), were overall significantly better for the stratified care condition than the stepped care condition. There was no significant difference in outcomes for complex cases, but standard cases had significantly better outcomes in the stratified care condition.

Outcome Measures

In line with NHS Talking Therapies service requirements, patients in the StratCare trial completed the PHQ-9 (Kroenke et al., 2001) and Generalised Anxiety Disorder-7 questionnaire (GAD-7; Spitzer et al., 2006) on a session-by-session basis. The Work and Social Adjustment Scale (WSAS; Mundt et al., 2002) was completed at the initial assessment and final session. The Standardised Assessment of Personality – abbreviate scale (SAPAS; Moran et al., 2003) was completed during the initial assessment session.

Primary Outcome Measure

The PHQ-9 is a measure of depression severity based on the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV) criteria for major depressive disorder (American Psychiatric Association, 1994; Kroenke et al., 2001; Appendix B). Patients rate the frequency of nine depression symptoms within the last two weeks on a 4-point Likert scale from 0 (‘not at all’) to 3 (‘nearly every day’), achieving an overall score of 0 to 27. Scores of 10 or over are rated as clinically significant (Kroenke et al., 2001). A change of six points or more is described as clinically significant change (Richards & Borglin, 2011). The psychometric properties of the PHQ-9 are good, with acceptable validity and internal reliability (Cronbach's α of 0.89; Cameron et al., 2008; Kroenke et al., 2001). The PHQ-9 has been validated for use in UK primary care services (Cameron et al., 2008).

Other Measures

The GAD-7 is a measure of anxiety severity based on the DSM-IV criteria for generalised anxiety disorder (American Psychiatric Association, 1994; Spitzer et al., 2006, Appendix C). Patients rate the frequency of seven anxiety symptoms within the last two weeks on a 4-point Likert scale from 0 (‘not at all’) to 3 (‘nearly every day’), achieving an overall score of 0 to 21. Scores of eight or over are considered indicative of anxiety disorders (Kroenke et al., 2007). A change of five points or more is described as clinically significant change (Richards & Borglin, 2011).

The WSAS measures impairment in functioning across five domains of daily activity (Mundt et al., 2002; Appendix D). Patients rate severity of impairment in functioning on a 9-point Likert scale from 0 (‘no impairment’) to 8 (‘very severe impairment’) in response to five statements, achieving an overall score of 0 to 40.

The SAPAS is a brief screening measure for personality disorders (Moran et al., 2003; Appendix E). Patients respond with ‘no’ (score of 0) or ‘yes’ (score of 1) to descriptive

statements, achieving an overall score of 0 to 8. Scores of three or over are considered indicative of likely presence of a personality disorder (Moran et al., 2003).

Ethical Approval

This specific secondary analysis received ethical approval from the University of Sheffield's Department of Psychology Research Ethics Committee (see Appendix F).

The StratCare trial received ethical approval from the West of Scotland Research Ethics Service (Reference 18/WS/0114; Appendix G) and approval from the Health Research Authority (Appendix H). This includes approval for future secondary data analysis. Interested patients and clinicians were provided with information on the trial. Participating clinicians provided written consent and participating patients provided verbal consent, thereby agreeing to future secondary data analysis. For further details see Delgadillo et al. (2022).

Original Dataset

The dataset contained self-reported information on the patient's age, gender, ethnicity, employment status and comorbidities. Additionally, information on referral source, case complexity, initially assigned treatment (LIT or HIT) and assigned therapist (anonymised) was obtained. Data of routinely completed outcome measures (PHQ-9, GAD-7, WSAS, SAPAS) was provided.

Within the dataset, patients were classified as 'standard' or 'complex' based on the cumulative case complexity model (Shippee et al., 2012). Complex cases represent patients who have poorer expected outcomes, compared to standard cases. A machine learning algorithm (LASSO regression analysis) was used to calculate a complexity score prior to treatment allocation. The following patient characteristics were included in the calculation: employment status, ethnic background, personality traits, functional impairment and level of depression and anxiety symptoms. These characteristics are independently associated with

poorer outcomes, but presence of multiple of these characteristics has a cumulative effect on outcomes (Delgadillo et al., 2017). Patients were classified as standard, or complex based on their complexity score, using an empirically derived cut-off score. Further details regarding the development of this machine learning approach can be found in Delgadillo et al. (2017).

Allocation to LIT or HIT followed either the stepped care model (*n* =583) or was based on a machine learning algorithm-based recommendation (*n* = 368). The computer algorithm allocated standard cases to LITs. Complex cases were allocated to HITs (Delgadillo et al., 2022). Patients receiving HIT were allocated to treatment modality, such as CBT of CfD, based on clinical judgement.

Data Selection

Pre-defined inclusion and exclusion criteria (see Table 1) were used to select cases for inclusion in the present study.

Table 1

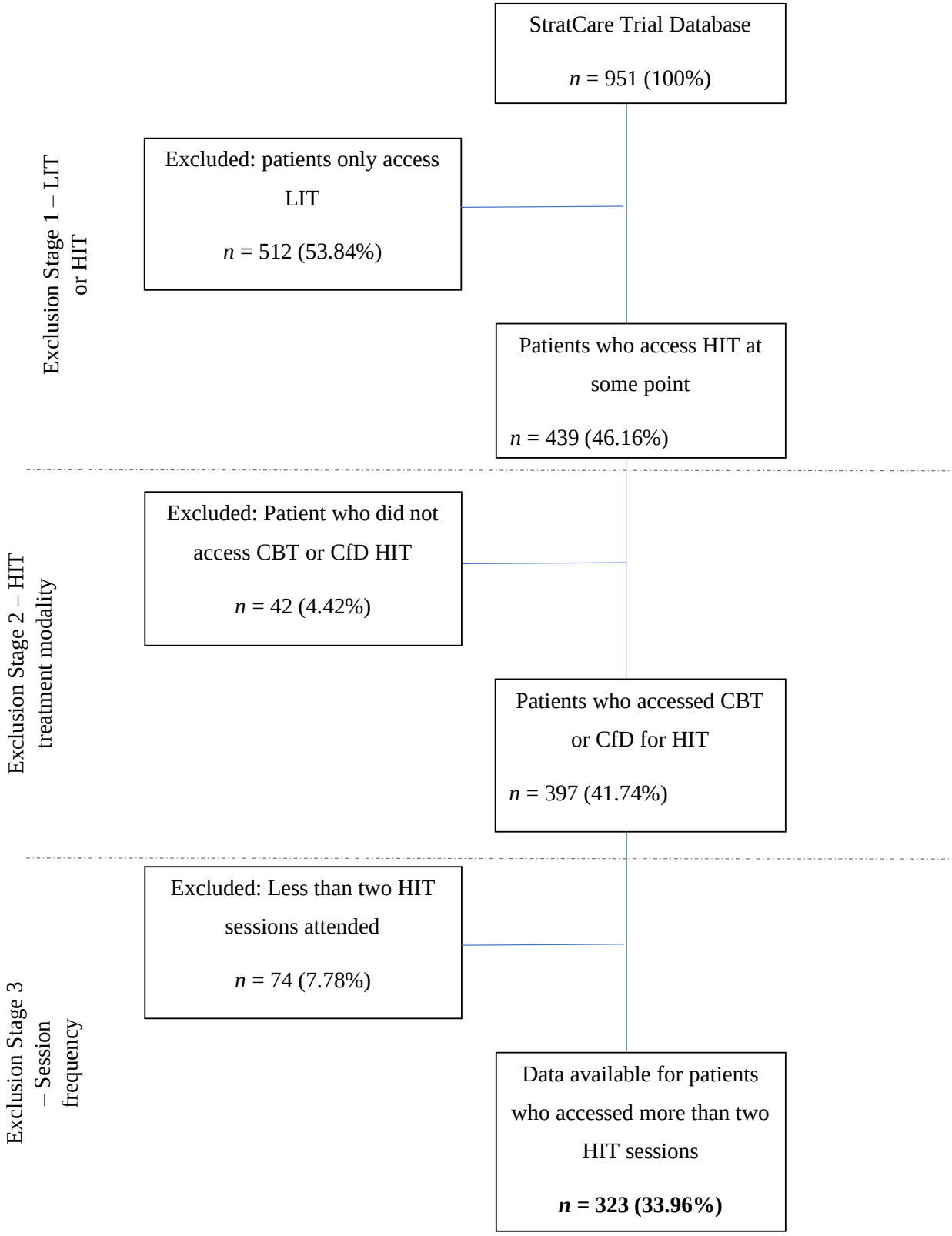
Overview of Patient Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Accessed high-intensity CBT or CfD.	Did not access high-intensity CBT or CfD.
Attended at least two intervention sessions.	Attended less than two intervention sessions.
Available case complexity classification.	Case complexity classification not available.

The original data set was screened against these criteria, as illustrated in Figure 1. Given the studies aims, the first step involved identifying all patients who accessed HITs. Following this, those who accessed high-intensity CBT or CfD were identified. To allow evaluation of treatment outcomes at least two HIT sessions had to be attended. As the PHQ-9 measures depression symptoms retrospectively for the last two weeks, availability of only the first HIT session PHQ-9 score would simply reflect pre-treatment symptom severity. At least one additional PHQ-9 measurement is required to assess symptom change during the course of treatment. Therefore, the third step involved excluding all patients who attended less than two HIT sessions. Following this, the identified sample (n = 323) met all set inclusion and exclusion criteria. As afore mentioned, patients accessed HITs either following the stepped care or stratified care approach. Consequently, the identified sample contained patients from both the stepped care and stratified care condition.

Figure 1

STROBE Diagram of Sample Selection from the StratCare Trial Database



Statistical Analysis

Data analysis was completed using IBM SPSS Statistics for Windows (version 28). Prior to data analysis, missing data was addressed using multiple imputation. This is because solely relying on complete case data can increase the risk of bias (Sterne et al., 2009). Depression, anxiety, and functional impairment scores were missing at HIT baseline (n = 40) and post-treatment (n = 1). Missing data was imputed using an expectation maximization method, with the following predictors: ethnicity, employment, SAPAS score, initial assessment PHQ-9, GAD-7, and WSAS score, treatment pathway. Subsequently, descriptive data analysis was completed, including a comparison of characteristics between the CBT and CfD samples using the Mann-Whitney U test for continuous variables and the Chi-Square test for categorical variables. Treatment effectiveness was established by calculating effect sizes in Excel using Minami's et al. (2008) formula. Effect sizes were interpreted in line with Cohen's (1988) guidance, with effect sizes classed as small (d = 0.2), medium (d = 0.5) and large (d = 0.8).

The primary analysis applied multiple regression to examine the relationship between case complexity and treatment response. The outcome (dependent) variable was post-treatment depression measured by the PHQ-9. The following independent variables were included in the model: treatment modality (CBT vs CfD), case complexity (standard vs complex case), interaction between treatment modality and case complexity, and baseline depression severity as measured by the PHQ-9. The main variable of interest was the interaction between treatment modality and case complexity. Baseline PHQ-9 at session 1 was included as a covariate to adjust for differences in depression severity at the start of treatment. Continuous variables (baseline PHQ-9 scores) were grand-mean centred to aid interpretation (Snijders et al., 2012). The study was sufficiently powered for the primary multiple regression analysis. Sample size calculations followed Cohen's (1992) table, with a multiple regression analysis with four independent variables at a significance level of $\alpha = .05$ and at 80% power, requiring a sample

size of 84 patients per treatment condition (CBT or CfD) for detecting a medium effect size. Expected effect size was informed by Delgadillo and Gonzalez Salas Duhne (2020), who found a medium effect size for differential treatment effects in their study.

Two different sensitivity analyses were completed to assess for robustness of results yielded by the multiple regression analysis. Firstly, a multi-level (mixed effects) regression model was used to assess for possible therapist effects on treatment outcomes. Therapist effects can account for differences in outcomes, therefore possible effects were examined in this study (Baldwin & Imel, 2013). The regression model included patients (level 1) nested within therapists (level 2). Patients with missing therapist data (e.g., unable to match patients to a specific therapist due to missig therapist identifier) were excluded from this analysis (n = 4). The regression model included all other variables specified in the primary analysis regression model. A random intercept for the therapist level was included in the model. Different sample size recommendations for examining therapist effects in multi-level modelling have been made in the literature, overall indicating samples sizes of around 1000 patients (Hox, 2010; Schiefele et al., 2017). This was not met in the given study, thus the sample was insufficiently powered and therapist effects were examined as a sensitivity analysis.

Secondly, the primary regression model described above was repeated in a case-control matched sample using propensity score matching (PSM). Due to the use of routinely collected clinical data, patients were not randomly allocated to either CBT or CfD, thus increasing risk of confounding by indication. PSM is a statistical technique that balances covariates across groups in a similar way as is achieved in randomised studies and can therefore help to minimise confounding bias (Rosenbaum & Rubin, 1983; Austin, 2011a). PSM utilises logistic regression to calculate a propensity score which predicts occurrences of a binary variable, in this instance HIT treatment allocation. A propensity score predicting allocation to CfD was calculated for each patient using the following predictor variables: age, gender, ethnicity, employment status

and initial assessment SAPAS, PHQ-9, GDAD and WSAS scores. Patients across both treatment conditions were then matched based on the propensity score using a one-to-one nearest neighbour approach, with exact matches prioritised. A conventional match tolerance (calliper of 0.2) was selected (Austin, 2011b). This resulted in a sample of 100 patients per treatment condition. Descriptive data analysis was completed for the PSM sample, including comparison of characteristics between the CBT and CfD group using an independent t-test or Mann-Whitney U test for continuous variables and the Chi-Square test of independence for categorical variables.

Secondary data analyses included reporting reliable improvement and reliable deterioration scores for both treatment conditions. Improvement and deterioration were determined based on changes in PHQ-9 scores between baseline and post-treatment. A reduction of six or more points is indicative of reliable improvement. An increase of six or more points is indicative of reliable deterioration (Richards & Borglin, 2011). This change in score meets Jacobson and Truax (1991) criteria for reliable change. Improvement and deterioration were determined based on changes in PHQ-9 scores between HIT baseline and post-treatment. A chi-square test was used to examine differences in deterioration rates and remission rates across both treatment conditions for complex cases. This analysis was also completed using the PSM sample.

Results

Selected Sample Characteristics

Out of the identified sample (*n* = 323), most patients accessed CBT (*n* = 223), were female (*n* = 215) and of a white British background (*n* = 302). Furthermore, most patients had a primary diagnosis of an affective and/or anxiety disorder. Approximately 80% of patients had a clinically significant PHQ-9 score of 10 or over. Over a third of patients were classified as complex (*n* = 129), with more complex cases accessing CBT than CfD. Further

characteristics are summarised in Table 2. Seventy-four therapists offered HITs to the included patients. Therapist information was not available for four patients. Sample characteristics were mostly similar across both HITs. Significantly higher rates of unemployment, complex cases, anxiety, and ‘other’ disorders were reported for the CBT sample. Significantly more patients with affective disorder accessed CfD.

Table 2

Patient Characteristics of Selected Sample

	Full Sample (n = 323)	CBT (n = 223)	CfD (n = 100)	Comparison of CBT and CfD sample	
Demographic/ Characteristic	Mean (SD)/%	Mean (SD)/%	Mean (SD)/%	Test statistics (d. f.)	p
Demographics					
Age (years)	39.6 (14.5)	38.6 (14.1)	41.85 (15.2)	U = 9803	.082
Female	215 (66.6%)	144 (64.6%)	71 (71%)	χ^2 (1) = 1.281	.258
White	302 (93.5%)	207 (92.8%)	95 (95%)	χ^2 (1) = .537	.464
Unemployed	89 (27.6%)	70 (31.4%)	19 (19%)	χ^2 (1) =5.309	.021*
Clinical characteristics					
Primary diagnosis					
Affective disorder	135 (41.8%)	66 (29.6%)	69 (69%)	χ^2 (1) = 44.064	< .001*
Anxiety disorder	98 (30.4%)	82 (37.7%)	16 (16%)	χ^2 (1) = 14.094	< .001*
Mixed (Affective and Anxiety disorder)	32 (9.9%)	25 (11.2%)	7 (7%)	χ^2 (1) = 1.371	.242
Other/ Missing	58 (18%)	50 (22.4%)	8 (8%)	χ^2 (1) = 12.812	< .001*
Prescribed medication	195 (60.4%)	136 (61%)	59 (59%)	χ^2 (1) = .451	.502
Comorbid long-term medical illness	51 (15.8%)	36 (16.1%)	15 (15%)	χ^2 (1) = .073	.787
Disability	39 (12.1%)	31 (13.9%)	8 (8%)	χ^2 (1) = 2.253	.133
SAPAS score (initial assessment)	4.28 (1.5)	4.35 (1.5)	4.12 (1.5)	U = 10175.5	.2
Complex Cases	129 (39.9%)	102 (45.7%)	27 (27%)	χ^2 (1) = 10.108	.001*
HIT session number	8.5 (5.3)	8.9 (5.8)	7.5 (4.3)	U = 10090	.171

Table 2 (continued)

	Full Sample (n = 323)	CBT (n = 223)	CfD (n = 100)	Comparison of CBT and CfD sample	
Demographic/ Characteristic	Mean (SD)/%	Mean (SD)/%	Mean (SD)/%	Test statistics (d. f.)	p
HIT baseline PHQ-9	14.94 (6.01)	14.88 (6.37)	15.07 (5.141)	U = 11045.5	.893
HIT baseline PHQ-9 score ≥ 10	261 (80.8%)	176 (78.924%)	85 (85%)	χ^2 (1) = 1.643	.2
HIT baseline GAD-7	13.76 (5.17)	13.96 (5.27)	13.31 (4.952)	U = 10100	.175
HIT baseline WSAS	21.76 (9.06)	21.66 (9.42)	22.1 (8.243)	U = 10972.5	.819
Final PHQ-9	8.72 (6.47)	9.11 (6.87)	7.87 (5.89)	U = 10315.5	.281
Final GAD-7	8.04 (5.77)	8.34 (6.01)	7.38 (5.154)	U = 10260.5	.251
Final WSAS	14.64 (10.17)	14.6 (10.63)	14.74 (9.097)	U = 10863.5	.712

Note. SD = Standard deviation, SAPAS = Standardized Assessment of Personality – Abbreviated Scale, PHQ-9 = Patient Health Questionnaire-9, GAD-7 = Generalised Anxiety Disorder-7, WSAS = Work and Social Adjustment Scale. U = Mann-Whitney U test, χ^2 = Chi-Square test of independence, * indicates statistically significant difference between CBT and CfD sample.

Treatment Efficacy

CBT and CfD were found to have large pre-post treatment effect sizes (d > 0.8), indicating that both treatments were effective in reducing PHQ-9 scores (see Table 3). Large effect sizes remained when examining standard and complex cases separately for each treatment modality. This suggests that CBT and CfD were effective treatments for both, standard and complex cases. Overall, PHQ-9 score decreased following HIT.

Table 3

Overview Effect Sizes

Sample	HIT baseline PHQ-9 (SD)	Final PHQ-9 (SD)	Effect Size (*d*)	95% CI
Standard Cases (*n* = 194)	12.81 (5.74)	7.1 (5.33)	0.99	0.81 - 1.17
Complex Cases (*n* = 129)	18.07 (4.85)	11.17 (7.24)	1.41	1.16 - 1.67
CBT (*n* = 223)	14.88 (6.37)	9.11 (6.87)	0.90	0.75 - 1.06
Standard (*n* = 121)	11.94 (6.01)	6.76 (5.47)	0.86	0.63 – 1.08
Complex (*n* = 102)	18.28 (4.79)	11.89 (7.34)	1.32	1.05 – 1.6
CfD (*n* = 100)	15.07 (5.14)	7.87 (5.89)	1.39	1.09 - 1.69
Standard (*n* = 73)	14.26 (4.96)	7.66 (5.08)	1.32	0.99 - 1.65
Complex (*n* = 27)	17.26 (5.06)	8.44 (6.23)	1.69	1.02 - 2.36

Note. SD = Standard Deviation; CI = Confidence Interval; *d* = Cohen's *d* effect size.

Primary Data Analysis

The results of the primary regression analysis are summarised in Table 4 and outlined regression coefficients that were included in the model. The analysis was powered to detect an expected medium effect size. As reported in Table 4, the regression analysis was indeed able to detect a medium effect size, R^2= 0.236 (Cohen, 1988). Out of the included predictor variables, only baseline PHQ-9 scores were found to be significant predictors of post-treatment PHQ-9 scores. Given a one-point increase in baseline PHQ-9 scores, a 0.421 unit change above the sample mean in post-intervention PHQ-9 scores would be expected (*SE* = .06, $p < .001$). This means that higher PHQ-9 baseline scores are significantly predictive of higher post-intervention PHQ-9 scores. As illustrated in Figure 2, post-intervention depression scores were higher for complex cases in the CBT condition, and lower for complex cases accessing CfD. However, treatment modality, case complexity and the interaction between case complexity and treatment modality were not significantly associated with post-intervention depression scores.

Table 4

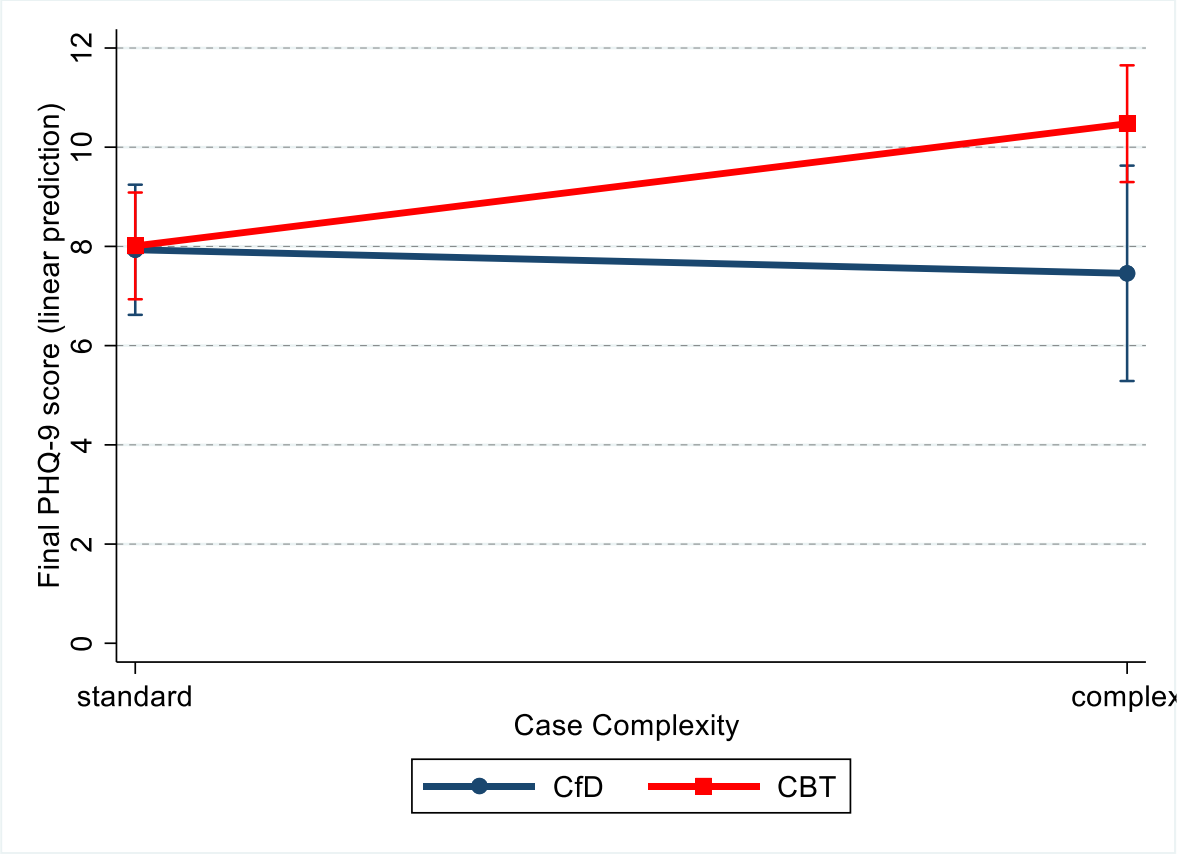
Fixed Coefficients of Multiple Linear Regression Analysis Predicting Post-Intervention PHQ-9 Scores (n = 323)

	R^2 = .236					
	B	SE	*t*	95% CI		*p*
Intercept	7.932	0.667	11.894	6.62	9.224	<.001
Treatment modality (CBT)	.078	.854	.091	-1.603	1.759	.927
Case complexity (Complex)	-.475	1.294	-.367	-3.02	2.07	.714
Baseline PHQ-9 score (mc)	.421	.06	7.052	0.303	0.538	<.001
Case complexity x treatment modality (Complex x CBT)	2.938	1.505	1.952	-.023	5.9	.052

Note. Baseline PHQ-9 scores (mc) = Mean-centred values; Treatment modality: 0 = CfD, 1 = CBT; Case complexity: 0 = standard case; 1 = complex case; Case complexity x Treatment modality: this interaction term is main variable of interest; B = Fixed regression coefficient; SE = Standard Error; CI = Confidence Interval.

Figure 2

Interaction Plot Illustrating Effect of Interaction of Case Complexity and Treatment Type including 95% Confidence Intervals (n = 323)



Sensitivity Analysis

Therapist effects. The first sensitivity analysis, a multi-level regression analysis, additionally modelled (i.e., controlled for) the possible impact of therapist effects on post-intervention PHQ-9 scores. Compared to the single-level model, a two-level regression model did not improve model fit (see Table 5). Although overall effect size reduced, it still remained in the medium range, R^2= 0.169 (Cohen, 1988) Like in the single-level model, only baseline PHQ-9 severity was a significant predictor of post-treatment depression scores, *B* = 0.421, *SE*

= .06, $p < .001$. Therapist effects did not contribute to the model, indicating that therapist effects did not predict post-treatment depression scores in this sample.

Table 5

Fixed Coefficients of Multi-Level Regression Model Predicting Post-Intervention PHQ-9 Scores and Controlling for Therapist Effects (n = 319; k = 74)

	R^2 = .169					
	B	SE	*t*	95% CI		*p*
Intercept	7.877	.674	11.679	6.55	9.204	<.001
Treatment modality (CBT)	.109	.866	.126	-1.585	1.813	.9
Case complexity (Complex)	-.420	1.302	0.323	-2.982	2.141	.747
Baseline PHQ-9 score (mc)	.421	.06	7.008	.303	.539	<.001
Case complexity x treatment modality (Complex x CBT)	2.907	1.517	1.917	-.077	5.891	.056

Note. Baseline PHQ-9 scores (mc) = Mean-centred values; Treatment modality: 0 = CfD, 1 = CBT; Case complexity: 0 = standard case; 1 = complex case; Case complexity x Treatment modality: this interaction term is main variable of interest; B = Fixed regression coefficient; SE = Standard Error; CI: Confidence Interval.

Propensity Score Matching. PSM was used to match all CfD patients (*n* = 100) to one CBT case with similar baseline features, yielding a total sample of 200 cases with balanced characteristics. Demographic and clinical characteristics of the PSM sample, as well as statistical comparison between the CfD and CBT sample, are summarised in Appendix I. Significant differences between CfD and CBT sample remained for affective disorder ($\chi 2$ (1) = 40.699, $p < .001$), anxiety disorder ($\chi 2$ (1) = 11.321, $p < .001$) and frequency of other disorders ($\chi 2$ (1) = 23.86, $p < .001$), with the latter two being more frequent in the CBT group. Additionally, significant differences appeared for the number of patients with a baseline PHQ-9 score of or equal 10 ($\chi 2$ (1) = 5.711, $p = .017$) and baseline WSAS scores ($t(198) = 2.304$, $p = .022$), with the former being more frequent in CfD patients and latter higher for CBT patients.

The results of the multiple regression analysis based on the propensity score matched sample are summarised in Table 6. As shown in the results, the same fixed coefficients are

included compared to the previous regression models. Again, only baseline PHQ-9 severity was found to be a significant predictor of post-intervention depression scores, $B = 0.377$, $SE = .075$, $p < .001$. The observed effect size falls within the medium range, $R^2 = .123$ (Cohen, 1988).

Table 6

Fixed Coefficients of Multiple Regression Analysis Predicting Post-Intervention PHQ-9 Scores using PSM Sample (n = 200)

	R^2 = .123					
	B	SE	*t*	95% CI		*p*
Intercept	7.904	.623	12.693	6.676	9.132	<.001
Treatment modality (CBT)	1.139	.977	1.165	-.789	3.066	.245
Case complexity (Complex)	-.322	1.216	-.283	-2.742	2.053	.777
Baseline PHQ-9 score (mc)	.377	.075	5.014	.229	.526	<.001
Case complexity x treatment modality (Complex x CBT)	2.214	1.623	1.364	-.986	5.414	.174

Note. Baseline PHQ-9 scores (mc) = Mean-centred values; Treatment modality: 0 = CfD, 1 = CBT; Case complexity: 0 = standard case; 1 = complex case; Case complexity x Treatment modality: this interaction term is main variable of interest; B = Fixed regression coefficient; SE = Standard Error; CI: Confidence Interval.

Secondary Analyses

On average, PHQ-9 scores changed by 6.19 points (*SD* = 6.515) between the first and final HIT session. Changes in PHQ-9 scores from first and final HIT session ranged from -11 to 25 points, with positive values indicating improvement and negative values indicating deterioration. Reliable improvement and deterioration rates for the CBT and CfD samples are summarised in Table 7.

Table 7

Overview of Reliable Improvement and Deterioration Rates Pre- and Post-Intervention as Measured by the PHQ-9

	Reliable Improvement *n* (%)	Reliable Deterioration *n* (%)
Full sample		
CfD *(n = 100)*	62 (62%)	2 (2%)
CBT *(n = 223)*	107 (48%)	5 (2.2%)
Total *(n = 323)*	169 (52.3%)	7 (2.2%)
Complex cases subsample		
CfD *(n = 27)*	18 (66.6%)	1 (3.7%)
CBT *(n = 102)*	54 (52.9%)	3 (2.9%)
Total *(n = 129)*	72 (55.8%)	4 (3.1%)

A Chi-Square test of independence showed that there was a significant difference in reliable improvement rates between the CBT and CfD sample, $\chi^2(1) = 5.438$, $p = 0.02$. Patients who accessed CfD were more likely to have achieved reliable improvement than those in the CBT sample. However, no significant difference in reliable improvement rates between the CfD and CBT samples was found for complex cases only, $\chi^2(1) = 1.631$, $p = .202$. This suggests reliable improvement rates for complex cases were similar across both, CfD and CBT.

No significant difference in reliable deterioration rates was found between the CBT and CfD samples, $\chi^2(1) = 0.019$, $p = 0.89$. This indicates that the likelihood of deterioration was similar across both treatment modalities. Statistical comparison of reliable deterioration rates was not completed for complex cases only, due to violation of Chi-Square test assumptions given a low number of cases that showed reliable deterioration.

Analysis of reliable improvement and deterioration rates using the PSM sample revealed similar results (Appendix J).

Discussion

This study aimed to explore whether complex cases may show differential treatment response when accessing different treatments for depression symptoms. The findings contribute towards a growing field of using statistical methods to optimise outcomes in depression through improved treatment allocation (Cohen & DeRubeis al., 2018; Kessler et al., 2016).

Main Findings

Firstly, it was hypothesised that case complexity would have a main effect on treatment outcomes. Secondly, complex cases who received CBT were hypothesised to have better outcomes compared to complex cases who received CfD. Neither hypothesis was supported by the results of this study. On the contrary, case complexity, treatment modality, as well as the interaction between case complexity and treatment modality did not have significant main effects on treatment outcomes as measured by the PHQ-9. Furthermore, reliable improvement and reliable deterioration rates for complex cases were comparable across both treatment modalities.

Whilst effect sizes of different psychotherapies for depression, including direct comparison of CBT and CfD, have been found to be similar (Pybis et al., 2017), differential treatment responses tend to be noted when analysing specific patient subgroups (Cuijpers et al., 2020; Delgadillo & Gonzalez Salas Duhne, 2020). However, this study found that complex cases did not show differential treatment responses. This suggests that complex cases respond similarly to high-intensity CBT and CfD in a primary care setting.

Contribution to the Evidence Base

Although in this study the main regression analysis did not find differential treatment outcomes for complex cases, significant differences in reliable improvement rates were reported. Reliable improvement was significantly more frequent for CfD when evaluating both,

standard and complex cases together. This significant difference remained even when controlling for the distribution of complex cases across both conditions, but the difference was not found when looking at complex cases only. This suggests that there may be differential treatment responses for standard cases only; in other words less complex cases may have better outcomes when accessing CfD. A possible explanation may be that standard cases are more likely to present with a lower number of factors affecting outcomes, and thus the influence of such factors on differential treatment outcomes becomes more apparent. For instance, a patient from an ethnic minority background would be classed as a standard case in the absence of other complicating factors. Thus, they would be predicted to respond better to CfD than CBT in line with previous research (Delgadillo & Gonzalez Salas Duhne, 2020).

With this study reporting similar outcomes for CBT and CfD, the chosen treatment modality itself appears to be of less importance when treating complex cases for depression. This can be viewed as consistent with the common factor models of psychotherapy, such as the contextual model, which argue that it is the shared features of different psychotherapies which lead to change (Wampold & Imel, 2015). However, there are other alternative explanations for why complex cases showed similar responses to CBT and CfD treatment in this study. Firstly, the selected sample will have included treatment non-completers given that only two HIT sessions had to be attended for participant inclusion. A review on dose-response effects in psychotherapy recommended that for reliable improvement to be observed, at least four sessions should be offered to mild to moderate cases and eight sessions to moderate to severe cases (Robinson et al., 2019). In this study, the average session frequency was 8.5 (*SD* = 5.3). Although this number is above the recommended eight sessions, several participants will have not accessed sufficient sessions for reliable improvements to be noted. Secondly, other research highlights how session frequency may impact on outcomes. Pybis et al. (2017) found CBT to be more effective when over eight sessions were accessed, whereas CfD was more effective in

comparison when less than eight sessions were accessed. However, this study did not examine the impact of session frequency on outcomes for complex cases.

Only baseline depression severity was found to be predictive of post-intervention PHQ-9 symptoms, with higher baseline severity predicting higher post-intervention depression scores. This is consistent with the wider literature, where baseline depression severity has been repeatedly shown to be associated with treatment outcomes (Driessen et al., 2010; Kessler et al., 2016). In relation to clinical practice this suggests that regardless of whether a patient is classified as standard or complex, level of baseline depression symptoms is more indicative of treatment outcome.

With a lack of trials reporting reliable deterioration rates, both reliable improvement and deterioration was examined in this study (Cuijpers et al., 2018). Overall, only 56% of complex cases showed reliable improvement and 3.1% reliable deterioration, which is comparable to the wider literature (Cuijpers et al., 2014; Cuijpers et al., 2018; Pybis et al., 2017). This leaves a high number of patients who did not appear to benefit from the offered psychological intervention when solely considering depression symptomatology. However, complex cases are defined as patients with expected poor prognosis. Therefore, the sample of complex cases will have included patients who are unlikely to respond to offered treatment. This is consistent with research into so called ‘non-responders’ which indicates that psychological therapies do not vary much in efficacy for this patient group (Gloster et al., 2020). This may offer an alternative explanation as to why this study did not find differential treatment responses for complex cases. However, such explanations are speculative in nature and further research is required.

Strengths and Limitations

The data used in this study originated from a randomised control trial, with patients randomised to different clinical pathways (i.e., stepped care vs. stratified care). However, in the present study patients were not randomly allocated to the specific HITs of interest (CBT vs. CfD). Hence, the internal validity of the study was affected due to lack of randomization to these treatment modalities. This was addressed through appropriate statistical methodology, PSM, however significant sample differences remained. Additionally, propensity score matching led to significantly more CfD patients presenting with clinically significant initial PHQ-9 scores than in the matched CBT sample. Although the chosen calliper threshold may have affected quality of matching, a narrower calliper threshold would have likely resulted in a reduced sample size at the cost of more optimal matching (Austin, 2011b). Optimal PSM also relies on the inclusion of all relevant baseline variables, with unmeasured confounders possible when using routine clinical data (Beal & Kupzyk, 2014). For example, experience of childhood maltreatment can affect depression outcomes but was not measured in this patient sample (Nanni et al., 2012).

As therapist effects are widely recognised to impact on outcomes (Baldwin & Imel, 2013; Johns et al., 2019) multi-level regression analysis was completed to determine whether they were present in this study. Although therapist effects were not found, it is important to note that the sample size was likely insufficiently powered to detect such effects (Hox, 2010; Schiefele et al., 2017). The sample included in the multilevel analysis involved 319 patients to 74 therapists, much lower than recommendations made in the literature for sample sizes of around 1000 patients to detect therapist effects (Hox, 2010; Schiefele et al., 2017). Despite these limitations, completed sensitivity analyses yielded similar results to the main regression model. Therefore, results into therapist effects should be viewed with caution.

A limitation of this study is the absence of long-term follow-up data. The lack of research on the long-term effects of depression treatments has been highlighted in the literature, with some evidence suggesting that effect sizes of psychological therapies increase long-term (McPherson & Hengartner, 2019). Furthermore, to enhance generalisability of findings and sample size, participants were not required to have a formal diagnosis of major depression disorder, or a clinically significant PHQ-9 score. Thus, individuals presenting with subclinical depression were included. Given that individuals with subclinical depression have an increased risk of developing major depressive disorder, evaluating treatment outcomes for this patients group is of clinical relevance (Lee et al., 2019).

Patient-reported data and self-report outcome measures were utilised in this study. To increase internal validity clinician-reported outcomes measures should be included. The findings of this study are obtained from two NHS Talking Therapies services in the north of England and may therefore not be generalisable to other primary health care settings in other countries. Furthermore, a study by Pybis et al. (2017) highlights that CBT and CfD outcomes in NHS Talking Therapy services vary between services. It is unclear how the services in this study compare NHS Talking Therapy services nationally, and whether this affects findings. A final limitation is in relation to treatment allocation, with participants allocated to high-intensity CBT or CfD based on clinical judgement. However, this study did not allow for evaluation of whether client preference affected treatment allocation. This is of relevance, given that client preference can affect outcomes as noted by a recent review (Swift et al., 2018).

Future Research

Research into differential treatment response of complex cases, as defined by the cumulative case complexity model, is in the early stages. Given the high number of patients who did not show reliable change following the intervention, future research is essential to increase understanding of how depression outcomes can be improved for those classified as

complex cases. It is advised that future research utilises larger sample sizes to enable detection of possible therapist effect and to determine whether findings vary between services. Furthermore, future research should consider the inclusion of long-term follow-up data. Although high-intensity psychological interventions may be of similar efficacy in the short-term, long-term differences may exist for complex cases. This study only evaluated two treatment modalities. However, NICE (2022) recommend a range of psychological approaches as first-line intervention (e.g., interpersonal psychotherapy, short-term psychodynamic psychotherapy). Further research may consider evaluating how complex cases respond to other psychological interventions. As afore mentioned, impact of client preference should also be considered in future studies.

Combining the findings of Delgadillo et al. (2017) and this study, complex cases are noted to respond well to high-intensity treatments, irrespective of which high-intensity treatment is offered. Additionally, effects size for complex cases were found to be large, and larger than for standard cases. This suggest that complex cases can very much benefit from accessing high-intensity treatment. Therefore, future research may want to focus on understanding the mechanisms and factors that make high-intensity more effective than low-intensity treatments for complex cases.

Clinical Implications

Both, CBT and CfD are found to be of similar efficacy in reducing depression symptoms in complex cases within a primary mental health setting. This is consistent with recently updated NICE guidance where CBT and counselling are both recommended first-line psychological interventions for depression (NICE, 2022). In terms of implications for clinical practice, NHS Talking Therapies practitioners should ensure complex cases are matched to high-intensity treatments to ensure optimum outcomes (Delgadillo et al., 2017). This could take place through utilising personalised machine learning algorithms as has been successfully

demonstrated by Delgadillo et al. (2017; 2022). However, the use of such approaches to inform decision-making in clinical practice requires careful consideration. For instance, patient and therapist perspectives on using such approaches should be explored further. In line with a patient centred approach to care, patients should also be made aware when treatment-related recommendations are made utilising statistical prediction approaches.

Both CBT and CfD are acceptable treatments for complex cases with depression symptoms. However, this study also highlighted that just under half of complex cases did not show reliable improvements in response to CBT or CfD. Therefore, clinicians need to be aware that amongst complex cases there may a large proportion of patients who show no treatment response and require further specialist intervention as outlined by NICE guidance (NICE, 2022).

Conclusion

Research into depression outcomes has focused on understanding patient characteristics and other factors that account for differential treatment responses, with the aim of optimising treatment allocation in mental health services. This study found that both CBT and CfD are equally effective high-intensity treatments for complex cases. However, a high number of complex cases do not respond to CBT and CfD in primary mental health services, therefore different approaches to intervention may need to be considered.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.).

American Psychological Association. (2020). Journal article reporting standards. In *Publication manual of the American Psychological Association* (7th ed.). American Psychological Association. https://doi.org/10.1037/0000165-000

Austin, P. C. (2011a). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Austin, P. C. (2011b). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical Statistics, 10*(2), 150–161. https://doi.org/10.1002/pst.433

Baldwin, S. A., & Imel, Z. E. (2013). Therapist effects: Findings and methods. In M. J. Lambert (Eds), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 258–297). Wiley.

Banyard, H., Behn, A. J., & Delgadillo, J. (2021). Personality disorders and their relation to treatment outcomes in cognitive behavioural therapy for depression: a systematic review and meta-analysis. *Cognitive Therapy and Research*, *45*, 561-576. https://doi.org/10.1007/s10608-021-10203-x

Beal, S. J., & Kupzyk, K. A. (2013). An introduction to propensity scores. *The Journal of Early Adolescence, 34*(1), 66–92. https://doi.org/10.1177/0272431613503215

Beck, A. T. (1979). *Cognitive therapy of depression.* Guilford Press.

Bell, I. A. N., & Mellor, D. (2009). Clinical judgements: research and practice. *Australian Psychologist, 44*(2), 112-121. https://doi.org/10.1080/00050060802550023

Bridges, S. (2016). Mental health problems. In Craig R, Fuller E, Mindell J (Eds), *Health survey for England 2014: Health, social care and lifestyle*. Health and Social Care Information Centre. www.hscic.gov.uk/pubs/hse2014

Buckman, J. E. J., Saunders, R., Stott, J., Arundell, L.-L. ., O'Driscoll, C., Davies, M. R., Eley, T. C., Hollon, S. D., Kendrick, T., Ambler, G., Cohen, Z. D., Watkins, E., Gilbody, S., Wiles, N., Kessler, D., Richards, D., Brabyn, S., Littlewood, E., DeRubeis, R. J., & Lewis, G. (2021a). Role of age, gender and marital status in prognosis for adults with depression: an individual patient data meta-analysis. *Epidemiology and Psychiatric Sciences*, *30*, e42. https://doi.org/ 10.1017/S2045796021000342

Buckman, J. E. J., Saunders, R., Cohen, Z. D., Barnett, P., Clarke, K., Ambler, G., DeRubeis, R. J., Gilbody, S., Hollon, S. D., Kendrick, T., Watkins, E., Wiles, N., Kessler, D., Richards, D., Sharp, D., Brabyn, S., Littlewood, E., Salisbury, C., White, I. R., & Lewis, G. (2021b). The contribution of depressive "disorder characteristics" to determinations of prognosis for adults with depression: an individual patient data meta-analysis. *Psychological Medicine*, 1–14. https://doi.org/10.1017/s0033291721001367

Buckman, J. E. J., Saunders, R., Stott, J., Cohen, Z. D., Arundell, L.-L., Eley, T. C., Hollon, S. D., Kendrick, T., Ambler, G., Watkins, E., Gilbody, S., Kessler, D., Wiles, N., Richards, D., Brabyn, S., Littlewood, E., DeRubeis, R. J., Lewis, G., & Pilling, S. (2022). Socioeconomic indicators of treatment prognosis for adults with depression. *JAMA Psychiatry*, *79*(5), 406. https://doi.org/10.1001/jamapsychiatry.2022.0100

Cameron, I. M., Crawford, J. R., Lawton, K., & Reid, I. C. (2008). Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. *The British Journal of General Practice : The Journal of the Royal College of General Practitioners*, *58*(546), 32–36. https://doi.org/10.3399/bjgp08X263794

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). L. Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, Z. D., & DeRubeis, R. J. (2018). Treatment selection in depression. *Annual Review of Clinical Psychology, 14*(1), 209-236. https://doi.org/10.1146/annurev-clinpsy-050817-084746

Cohen, Z. D., & DeRubeis, R. J. (2018). Treatment selection in depression. *Annual Review of Clinical Psychology*, *14*(1), 209–236. https://doi.org/10.1146/annurev-clinpsy-050817-084746

Constantino, M. J., Boswell, J. F., Coyne, A. E., Swales, T. P., & Kraus, D. R. (2021). Effect of matching therapists to patients vs assignment as usual on adult psychotherapy outcomes: a randomized clinical trial. *JAMA Psychiatry*, *78*(9), 960-969 10.1001/jamapsychiatry. 2021.1221

Cuijpers, P., Karyotaki, E., de Wit, L., & Ebert, D. D. (2020). The effects of fifteen evidence-supported therapies for adult depression: a meta-analytic review. *Psychotherapy Research*, *30*(3), 279-293. https://doi.org/10.1080/10503307.2019.1649732

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2014). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: a meta-analysis. *Journal of Affective Disorders*, *159*, 118-126. https://doi.org/10.1016/j.jad.2014.02.026

Cuijpers, P., Reijnders, M., Karyotaki, E., de Wit, L., & Ebert, D. D. (2018). Negative effects of psychotherapies for adult depression: a meta-analysis of deterioration rates. *Journal of Affective Disorders*, *239*, 138-145. https://doi.org/10.1016/j.jad.2018.05.050

de Jonge, P., Huyse, F. J., Slaets, J. P. J., Söllner, W., & Stiefel, F. C. (2005). Operationalization of biopsychosocial case complexity in general health care: the INTERMED project. *Australian & New Zealand Journal of Psychiatry*, *39*(9), 795-799. https://doi.org/10.1080/j.1440-1614.2005.01684.x

Delgadillo, J., & Gonzalez Salas Duhne, P. (2020). Targeted prescription of cognitive–behavioral therapy versus person-centered counselling for depression using a machine learning approach. *Journal of Consulting and Clinical Psychology*, *88*(1), 14-24. https://doi.org/10.1037/ccp0000476

Delgadillo, J., Ali, S., Fleck, K., Agnew, C., Southgate, A., Parkhouse, L., Cohen, Z. D., DeRubeis, R. J., & Barkham, M. (2022). Stratified care vs stepped care for depression. *JAMA Psychiatry*, *79*(2), 101-108 https://doi.org/10.1001/jamapsychiatry.2021.3539

Delgadillo, J., Dawson, A., Gilbody, S., & Böhnke, J. R. (2017). Impact of long-term medical conditions on the outcomes of psychological therapy for depression and anxiety. *British Journal of Psychiatry*, *210*(1), 47–53. https://doi.org/10.1192/bjp.bp.116.189027

Delgadillo, J., Huey, D., Bennett, H., & McMillan, D. (2017). Case complexity as a guide for psychological treatment selection. *Journal of Consulting and Clinical Psychology*, *85*(9), 835-853. https://doi.org/10.1037/ccp0000231

Driessen, E., Cuijpers, P., Hollon, S. D., & Dekker, J. J. M. (2010). Does pretreatment severity moderate the efficacy of psychological treatment of adult outpatient depression? A meta-analysis. *Journal of Consulting and Clinical Psychology*, *78*(5), 668-680. https://doi.org/10.1037/a0020570

Elliott, R. (2004). *Learning emotion-focused therapy : the process-experiential approach to change*. American Psychological Association.

Finegan, M., Firth, N., Wojnarowski, C., & Delgadillo, J. (2018). Associations between socioeconomic status and psychological therapy outcomes: a systematic review and meta-analysis. *Depression and Anxiety*, *35*(6), 560–573. https://doi.org/10.1002/da.22765

Gloster, A. T., Rinner, M. T. B., Ioannou, M., Villanueva, J., Block, V. J., Ferrari, G., Benoy, C., Bader, K., & Karekla, M. (2020). Treating treatment non-responders: a meta-analysis of randomized controlled psychotherapy trials. *Clinical Psychology Review, 75*, 101810. https://doi.org/10.1016/j.cpr.2019.101810

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: a meta-analysis. *Psychological Assessment, 12*(1), 19-30. https://doi.org/10.1037/1040-3590.12.1.19

Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology*, *61*(2), 155-163. https://doi.org/10.1002/jclp.20108

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical psychology, 9*(3), 329-343. https://doi.org/10.1093/clipsy.9.3.329

Health Education England (2022). *National curriculum for high intensity cognitive behavioural therapy courses (fourth edition)*. https://www.hee.nhs.uk/sites/default/files/documents/HiT%20Curriculum%20%28name%20of%20document%20on%20HEE%20website%20says%20PWP%20Curriculum%204th%20Edition%202022_10%20Nov%202022%29.pdf

Hill, A. (2011). *Curriculum for counselling for depression: Continuing professional development for qualified therapists delivering high intensity interventions*. National IAPT Programme Team. https://webarchive.nationalarchives.gov.uk/20160302160209/http://www.iapt.nhs.uk/silo/files/curriculum-for-counselling-for-depression.pdf

Hox, J. (2010). *Multilevel analysis: techniques and applications* (2. Aufl.). England: Routledge.

IBM Corp. (2022). IBM SPSS Statistics for Windows (28.0). [Computer software].

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. https://doi.org/10.1037/0022-006x.59.1.12

Johns, R. G., Barkham, M., Kellett, S., & Saxon, D. (2019). A systematic review of therapist effects: a critical narrative update and refinement to review. *Clinical Psychology Review, 67*, 78–93. https://doi.org/10.1016/j.cpr.2018.08.004

Kessler, R. C., van Loo, H. M., Wardenaar, K. J., Bossarte, R. M., Brenner, L. A., Ebert, D. D., de Jonge, P., Nierenberg, A. A., Rosellini, A. J., Sampson, N. A., Schoevers, R. A., Wilcox, M. A., & Zaslavsky, A. M. (2016). Using patient self-reports to study heterogeneity of treatment effects in major depressive disorder. *Epidemiology and Psychiatric Sciences, 26*(1), 22–36. https://doi.org/10.1017/s2045796016000020

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine, 16*(9), 606-613. https://doi.org/10.1046/j.1525-1497.2001.16009606.x

Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Annals*

of Internal Medicine, 146(5), 317. https://doi.org/10.7326/0003-4819-146-5-200703060-00004

Lee, Y. Y., Stockings, E. A., Harris, M. G., Doi, S. A. R., Page, I. S., Davidson, S. K., & Barendregt, J. J. (2018). The risk of developing major depression among individuals with subthreshold depression: a systematic review and meta-analysis of longitudinal cohort studies. *Psychological Medicine, 49*(1), 92–102. https://doi.org/10.1017/s0033291718000557

Lorenzo-Luaces, L., DeRubeis, R. J., van Straten, A., & Tiemens, B. (2017). A prognostic index (PI) as a moderator of outcomes in the treatment of depression: a proof of concept combining multiple variables to inform risk-stratified stepped care models. *Journal of Affective Disorders, 213*, 78-85. https://doi.org/10.1016/j.jad.2017.02.010

Martell, C. R, Addis, M. E., & Jacobson, N. S. (2001). *Depression in context : Strategies for guided action*. W. W. Norton & Co.

McCrone, P., Dhanasiri, S., Patel, A., Knapp, M., & Lawton-Smith, S. (2008*). Paying the price: The cost of mental health care in England to 2026*. The Kings Fund. https://www.kingsfund.org.uk/publications/paying-price

McPherson, S., & Hengartner, M. P. (2019). Long-term outcomes of trials in the National Institute for Health and Care Excellence depression guideline. *BJPsych Open, 5*(5). https://doi.org/10.1192/bjo.2019.65

Mearns, D., Thorne, B., & Mcleod, J. (2013). *Person-centred counselling in action* (4th ed.). Sage.

Minami, T., Serlin, R. C., Wampold, B. E., Kircher, J. C., & Brown, G. J. (2008). Using clinical trials to benchmark effects produced in clinical practice. *Quality and Quantity*, *42*(4), 513.

Moran, P., Leese, M., Lee, T., Walters, P., Thornicroft, G., & Mann, A. (2003). Standardised assessment of personality – abbreviated scale (SAPAS): preliminary validation of a brief screen for personality disorder. *British Journal of Psychiatry, 183*(3), 228–232. https://doi.org/10.1192/bjp.183.3.228

Mulder, R. T. (2002). Personality pathology and treatment outcome in major depression: a review. *American Journal of Psychiatry, 159*(3), 359-371. https://doi.org/10.1176/appi.ajp.159.3.359

Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The work and social adjustment scale: a simple measure of impairment in functioning. *British Journal of Psychiatry, 180*(5), 461–464. https://doi.org/10.1192/bjp.180.5.461

Nanni, V., Uher, R., & Danese, A. (2012). Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: a meta-analysis. *American Journal of Psychiatry, 169*(2), 141-151. https://doi.org/10.1176/appi.ajp.2011.11020335

National Collaborating Centre for Mental Health (2023). *The improving access to psychological therapies manual*. (Version 6). https://www.england.nhs.uk/wp-content/uploads/2018/06/the-nhs-talking-therapies-manual-v6.pdf

National Institute for Health and Care Excellence (NICE, 2009). *Depression in adults with a chronic physical health problem: recognition and management*. [NICE Guideline No. 91]. https://www.nice.org.uk/guidance/cg91

National Institute for Health and Care Excellence (NICE, 2022). *Depression in adults: recognition and management*. [NICE Guideline No. 222]. https://www.nice.org.uk/guidance/ng222

Newton-Howes, G., Tyrer, P., Johnson, T., Mulder, R., Kool, S., Dekker, J., & Schoevers, R. (2014). Influence of personality on the outcome of treatment in depression: systematic review and meta-analysis. *Journal of Personality Disorders*, *28*(4), 577–593. https://doi.org/ 10.1521/pedi201327_070

Oliveira, C. A. de, Weber, B., dos Santos, J. L. F., Zucoloto, M. L., de Camargo, L. L., Zanetti, A. C. G., Rzewuska, M., & de Azevedo-Marques, J. M. (2022). Health complexity assessment in primary care: a validity and feasibility study of the INTERMED tool. *PLOS ONE*, *17*(2), e0263702. https://doi.org/10.1371/journal.pone.0263702

Pybis, J., Saxon, D., Hill, A., & Barkham, M. (2017). The comparative effectiveness and efficiency of cognitive behaviour therapy and generic counselling in the treatment of depression: evidence from the 2nd UK National Audit of psychological therapies. *BMC Psychiatry*, *17*(215). https://doi.org/10.1186/s12888-017-1370-7

Richards, D. A., & Borglin, G. (2011). Implementation of psychological therapies for anxiety and depression in routine practice: two-year prospective cohort study. *Journal of Affective Disorders, 133*(1-2), 51–60. https://doi.org/10.1016/j.jad.2011.03.024

Robinson, L., Delgadillo, J., & Kellett, S. (2019). The dose-response effect in routinely delivered psychological therapies: a systematic review. *Psychotherapy Research, 30*(1), 79–96. https://doi.org/10.1080/10503307.2019.1566676

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55. https://doi.org/10.1093/biomet/70.1.41

Roth, A., & Pilling, S. (2008). Using an evidence-based methodology to identify the competences required to deliver effective cognitive and behavioural therapy for

depression and anxiety disorders. *Behavioural and Cognitive Psychotherapy, 36*(2), 129-147. https://doi.org/10.1017/s1352465808004141

Ruscio, A. M., & Holohan, D. R. (2006). Applying empirically supported treatments to complex cases: ethical, empirical, and practical considerations. *Clinical Psychology: Science and Practice, 13*(2), 146-162. https://doi.org/https://doi.org/10.1111/j.1468-2850.2006.00017.x

Sanders, P., & Hill, A. (2014). *Counselling for depression: A person-centred and experiential approach to practice*. Sage.

Schiefele, A.-K., Lutz, W., Barkham, M., Rubel, J., Böhnke, J., Delgadillo, J., Kopta, M., Schulte, D., Saxon, D., Nielsen, S. L., & Lambert, M. J. (2017). Reliability of therapist effects in practice-based psychotherapy research: a guide for the planning of future studies. *Administration and Policy in Mental Health and Mental Health Services Research, 44*(5), 598–613. https://doi.org/10.1007/s10488-016-0736-3

Shippee, N. D., Shah, N. D., May, C. R., Mair, F. S., & Montori, V. M. (2012). Cumulative complexity: a functional, patient-centered model of patient complexity can improve research and practice. *Journal of Clinical Epidemiology, 65*(10), 1041-1051. https://doi.org/10.1016/j.jclinepi.2012.05.005

Snijders, T.A.B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder. *Archives of Internal Medicine, 166*(10), 1092. https://doi.org/10.1001/archinte.166.10.1092

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ, 338*, b2393. https://doi.org/10.1136/ bmj.b2393

Swift, J. K., Callahan, J. L., Cooper, M., & Parkin, S. R. (2018). The impact of accommodating client preference in psychotherapy: a meta-analysis. *Journal of Clinical Psychology, 74*(11), 1924– 1937. https://doi.org/10.1002/jclp.22680

van Bronswijk, S. C., Lemmens, L. H. J. M., Huibers, M. J. H., & Peeters, F. P. M. L. (2021). Selecting the optimal treatment for a depressed individual: clinical judgment or statistical prediction? *Journal of Affective Disorders, 279*, 149-157. https://doi.org/10.1016/j.jad.2020.09.135

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate :the evidence for what makes psychotherapy work* (Second edition.). Routledge.

Appendices

Appendix A

The Journal Article Reporting Standards for quantitative studies (APA, 2020)

JARS–**Quant** | Table 1
Information Recommended for Inclusion in Manuscripts
That Report New Data Collections Regardless of Research Design

Title and Title Page

Title

- Identify main variables and theoretical issues under investigation and the relationships between them.
- Identify the populations studied.

Author Note

- Provide acknowledgment and explanation of any special circumstances, including
  - registration information if the study has been registered
  - use of data also appearing in previous publications
  - prior reporting of the fundamental data in dissertations or conference papers
  - sources of funding or other support
  - relationships or affiliations that may be perceived as conflicts of interest
  - previous (or current) affiliation of authors if different from location where the study was conducted
  - contact information for the corresponding author
  - additional information of importance to the reader that may not be appropriately included in other sections of the paper

Abstract

Objectives

- State the problem under investigation, including main hypotheses.

Participants

- Describe subjects (nonhuman animal research) or participants (human research), specifying their pertinent characteristics for the study; in animal research, include genus and species. Participants are described in greater detail in the body of the paper.

Study Method

- Describe the study method, including
  - research design (e.g., experiment, observational study)
  - sample size
  - materials used (e.g., instruments, apparatus)
  - outcome measures
  - data-gathering procedures, including a brief description of the source of any secondary data. If the study is a secondary data analysis, so indicate.

Findings

- Report findings, including effect sizes and confidence intervals or statistical significance levels.

Conclusions

- State conclusions, beyond just results, and report the implications or applications.

Introduction

Problem

- State the importance of the problem, including theoretical or practical implications.

Review of Relevant Scholarship

- Provide a succinct review of relevant scholarship, including
  - relation to previous work
  - differences between the current report and earlier reports if some aspects of this study have been reported on previously

Hypothesis, Aims, and Objectives

- State specific hypotheses, aims, and objectives, including
  - theories or other means used to derive hypotheses
  - primary and secondary hypotheses
  - other planned analyses
- State how hypotheses and research design relate to one another.

Method

Inclusion and Exclusion

- Report inclusion and exclusion criteria, including any restrictions based on demographic characteristics.

Participant Characteristics

- Report major demographic characteristics (e.g., age, sex, ethnicity, socioeconomic status) and important topic-specific characteristics (e.g., achievement level in studies of educational interventions).
- In the case of animal research, report the genus, species, and strain number or other specific identification, such as the name and location of the supplier and the stock designation. Give the number of animals and the animals' sex, age, weight, physiological condition, genetic modification status, genotype, health–immune status, drug or test naïveté, and previous procedures to which the animal may have been subjected.

Sampling Procedures

- Describe procedures for selecting participants, including
  - sampling method if a systematic sampling plan was implemented
  - percentage of sample approached that actually participated
  - whether self-selection into the study occurred (either by individuals or by units, such as schools or clinics)
- Describe settings and locations where data were collected as well as dates of data collection.
- Describe agreements and payments made to participants.
- Describe institutional review board agreements, ethical standards met, and safety monitoring.

Sample Size, Power, and Precision

- Describe the sample size, power, and precision, including
  - intended sample size
  - achieved sample size, if different from the intended sample size
  - determination of sample size, including
    - › power analysis, or methods used to determine precision of parameter estimates
    - › explanation of any interim analyses and stopping rules employed

Measures and Covariates

- Define all primary and secondary measures and covariates, including measures collected but not included in the report.

Data Collection

- Describe methods used to collect data.

Quality of Measurements

- Describe methods used to enhance the quality of measurements, including
  - training and reliability of data collectors
  - use of multiple observations

Instrumentation

- Provide information on validated or ad hoc instruments created for individual studies, for individual studies (e.g., psychometric and biometric properties).

Masking

- Report whether participants, those administering the experimental manipulations, and those assessing the outcomes were aware of condition assignments.
- If masking took place, provide a statement regarding how it was accomplished and whether and how the success of masking was evaluated.

Psychometrics

- Estimate and report values of reliability coefficients for the scores analyzed (i.e., the researcher's sample), if possible. Provide estimates of convergent and discriminant validity where relevant.
- Report estimates related to the reliability of measures, including
  - interrater reliability for subjectively scored measures and ratings
  - test–retest coefficients in longitudinal studies in which the retest interval corresponds to the measurement schedule used in the study
  - internal consistency coefficients for composite scales in which these indices are appropriate for understanding the nature of the instruments being used in the study
- Report the basic demographic characteristics of other samples if reporting reliability or validity coefficients from those samples, such as those described in test manuals or in norming information for the instrument.

Conditions and Design

- State whether conditions were manipulated or naturally observed. Report the type of design as per the JARS–Quant tables:
  - experimental manipulation with participants randomized
    - › Table 2 and Module A
  - experimental manipulation without randomization
    - › Table 2 and Module B
  - clinical trial with randomization
    - › Table 2 and Modules A and C
  - clinical trial without randomization
    - › Table 2 and Modules B and C
  - nonexperimental design (i.e., no experimental manipulation): observational design, epidemiological design, natural history, and so forth (single-group designs or multiple-group comparisons)
    - › Table 3
  - longitudinal design
    - › Table 4
  - *N*-of-1 studies
    - › Table 5
  - replications
    - › Table 6
- Report the common name given to designs not currently covered in JARS–Quant.

Data Diagnostics

- Describe planned data diagnostics, including
  - criteria for post-data-collection exclusion of participants, if any
  - criteria for deciding when to infer missing data and methods used for imputation of missing data
  - definition and processing of statistical outliers
  - analyses of data distributions
  - data transformations to be used, if any

Analytic Strategy

- Describe the analytic strategy for inferential statistics and protection against experiment-wise error for
  - primary hypotheses
  - secondary hypotheses
  - exploratory hypotheses

Results

Participant Flow

- Report the flow of participants, including
  - total number of participants in each group at each stage of the study
  - flow of participants through each stage of the study (include figure depicting flow, when possible; see the JARS–Quant Participant Flowchart)

Recruitment

- Provide dates defining the periods of recruitment and repeated measures or follow-up.

Statistics and Data Analysis

- Provide information detailing the statistical and data-analytic methods used, including
  - missing data
    - › frequency or percentages of missing data
    - › empirical evidence and/or theoretical arguments for the causes of data that are missing—for example, missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR)
    - › methods actually used for addressing missing data, if any
  - descriptions of each primary and secondary outcome, including the total sample and each subgroup, that includes the number of cases, cell means, standard deviations, and other measures that characterize the data used
  - inferential statistics, including
    - › results of all inferential tests conducted, including exact *p* values if null hypothesis significance testing (NHST) methods were used, and reporting the minimally sufficient set of statistics (e.g., *df*s, mean square [*MS*] effect, *MS* error) needed to construct the tests
    - › effect-size estimates and confidence intervals on estimates that correspond to each inferential test conducted, when possible
    - › clear differentiation between primary hypotheses and their tests–estimates, secondary hypotheses and their tests–estimates, and exploratory hypotheses and their test–estimates

Statistics and Data Analysis (*continued*)

  - complex data analyses—for example, structural equation modeling analyses (see also Table 7), hierarchical linear models, factor analysis, multivariate analyses, and so forth, including
    - › details of the models estimated
    - › associated variance–covariance (or correlation) matrix or matrices
    - › identification of the statistical software used to run the analyses (e.g., SAS PROC GLM or the particular R package)
  - estimation problems (e.g., failure to converge, bad solution spaces), regression diagnostics, or analytic anomalies that were detected and solutions to those problems.
  - other data analyses performed, including adjusted analyses, if performed, indicating those that were planned and those that were not planned (though not necessarily in the level of detail of primary analyses).
- Report any problems with statistical assumptions and/or data distributions that could affect the validity of findings.

Discussion

Support of Original Hypotheses

- Provide a statement of support or nonsupport for all hypotheses, whether primary or secondary, including
  - distinction by primary and secondary hypotheses
  - discussion of the implications of exploratory analyses in terms of both substantive findings and error rates that may be uncontrolled

Similarity of Results

- Discuss similarities and differences between reported results and work of others.

Interpretation

- Provide an interpretation of the results, taking into account
  - sources of potential bias and threats to internal and statistical validity
  - imprecision of measurement protocols
  - overall number of tests or overlap among tests
  - adequacy of sample sizes and sampling validity

Generalizability

- Discuss generalizability (external validity) of the findings, taking into account
  - target population (sampling validity)
  - other contextual issues (setting, measurement, time; ecological validity)

Implications

- Discuss implications for future research, program, or policy.

Appendix B

The Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001)

Over the *last two weeks*, how often have you been bothered by any of the following problems?

	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or taking asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself – or that you are a failure of have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed? Or the opposite – being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead or of hurting yourself in some way	0	1	2	3

If you checked off *any* problems, how *difficult* have these problems made it for you to do your work, take care of things at home, or get along with other people?

Not difficult at all	Somewhat difficult	Very difficult	Extremely difficult
…	…	…	…

Appendix C

Generalised Anxiety Disorder-7 (GAD-7; Spitzer et al., 2006)

GAD-7

Over the last 2 weeks, how often have you been bothered by the following problems?	Not at all	Several days	More than half the days	Nearly every day
1. Feeling nervous, anxious or on edge	0	1	2	3
2. Not being able to stop or control worrying	0	1	2	3
3. Worrying too much about different things	0	1	2	3
4. Trouble relaxing	0	1	2	3
5. Being so restless that it is hard to sit still	0	1	2	3
6. Becoming easily annoyed or irritable	0	1	2	3
7. Feeling afraid as if something awful might happen	0	1	2	3

Total Score ___ = Add Columns ___ + ___ + ___

If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?

Not difficult at all	Somewhat difficult	Very difficult	Extremely difficult
☐	☐	☐	☐

Appendix D

The Work and Social Adjustment Scale (WSAS; Mundt et al., 2002)

Rate each of the following questions on a 0 to 8 scale: 0 indicates no impairment at all and 8 indicates very severe impairment.

	Not at							Very Severely
1. Because of my [disorder], my ability to work is impaired. '0' means not at all impaired, '8' means very severely impaired to the point I can't work.	0	1	2	3	4	5	6	7	8
2. Because of my [disorder], my home management (cleaning, tidying, shopping, cooking, looking after home or children, paying bills) is impaired.	0	1	2	3	4	5	6	7	8
3. Because of my [disorder], my social leisure activities (with other people, such as parties, bars, clubs, outings, visits, dating, home entertainment) are impaired.	0	1	2	3	4	5	6	7	8
4. Because of my [disorder] my private leisure activities (done alone, such as reading, gardening, collecting, sewing, walking alone) are impaired.	0	1	2	3	4	5	6	7	8
5. Because of my [disorder] my ability to form and maintain close relationships with others, including those I live with, is impaired.	0	1	2	3	4	5	6	7	8

Appendix E

The Standardised Assessment of Personality – Abbreviated Scale (Moran et al., 2003)

Please ask your patients the following questions. Only circle a response if the patient thinks that the description applies most of the time and in most situations.

1.	In general, do you have difficulty making and keeping friends?	YES	NO
2.	Would you normally describe yourself as a loner?	YES	NO
3.	In general, do you trust other people?	YES	NO
4.	Do you normally lose your temper easily?	YES	NO
5.	Are you normally an impulsive sort of person?	YES	NO
6.	Are you normally a worrier?	YES	NO
7.	In general, do you depend on others a lot?	YES	NO
8.	In general, are you a perfectionist?	YES	NO

Appendix F

Ethical Approval for Research Study

The University Of Sheffield.

Downloaded: 09/01/2022
Approved: 15/12/2021

[redacted]
Registration number: [redacted]
Psychology
Programme: Doctorate in Clinical Psychology

Dear [redacted]

PROJECT TITLE: Treatment of complex cases of depression in Improving Access to Psychological Therapies services: a quantitative study.
APPLICATION: Reference Number 044533

This letter confirms that you have signed a University Research Ethics Committee-approved self-declaration to confirm that your research will involve only existing research, clinical or other data that has been robustly anonymised. You have judged it to be unlikely that this project would cause offence to those who originally provided the data, should they become aware of it.

As such, on behalf of the University Research Ethics Committee, I can confirm that your project can go ahead on the basis of this self-declaration.

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since full ethical review may be required.

Yours sincerely

Department Of Psychology Research Ethics Committee
Departmental Ethics Administrator

Appendix G

Evidence of Ethical Approval for StratCare trial by West Scotland Research Ethics Service

WoSRES
West of Scotland Research Ethics Service

NHS
Greater Glasgow and Clyde

Dr Jaime Delgadillo
Lecturer in Clinical Psychology
University of Sheffield
Clinical Psychology Unit
University of Sheffield
Cathedral Court, Floor F
1 Vicar Lane, Sheffield
S1 1HD

West of Scotland REC 5
West of Scotland Research Ethics Service
West Glasgow Ambulatory Care Hospital
Dalnair Street
Glasgow
G3 8SJ

Date 18 July 2018

Direct line 0141 232 1809
E-mail WoSREC5@ggc.scot.nhs.uk

Please note: This is the favourable opinion of the REC only and does not allow you to start your study at NHS sites in England until you receive HRA Approval

Dear Dr Delgadillo

Study title:	**Pragmatic randomised controlled trial of a stratified care model for depression and anxiety**
REC reference:	**18/WS/0114**
Protocol number:	**152958**
IRAS project ID:	**247945**

Thank you for your letter of 11 July 2018, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to make a request to postpone publication, please contact hra.studyregistration@nhs.net outlining the reasons for your request.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission must be obtained from each host organisation prior to the start of the

study at the site concerned.

Management permission should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA and HCRW Approval (England and Wales)/ NHS permission for research is available in the Integrated Research Application System, at www.hra.nhs.uk or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database within 6 weeks of recruitment of the first participant (for medical device studies, within the timeline determined by the current registration and publication trees).

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from the HRA. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

NHS sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Covering letter on headed paper [Cover letter]		25 May 2018
Covering letter on headed paper [Cover letter]		08 June 2018
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [StratCare Insurance Certificate]		25 May 2018
Instructions for use of medical device [StratCare App: Instructions for use of non-medical device (Simply stated as MAY18)]	1	01 May 2018
IRAS Application Form [IRAS_Form_01062018]		01 June 2018
Letter from funder [Funding letter]		27 May 2018
Letter from sponsor [Sponsorship confirmation letter]		23 May 2018
Letter from statistician [Statistical review]		19 May 2018
Other [HRA definition of NON-MEDICAL device]		
Other [StratCare_Patients_Consent_process_v1]	1	04 July 2018
Other [StratCare_Patients_Info_Sheet_v1]	1	04 July 2018
Other [StratCare_Therapists_Consent_form_v3]	3	04 July 2018
Other [StratCare_Trial_Protocol_v2]	2	11 July 2018
Other [Cover letter 11.07.18]		11 July 2018
Participant information sheet (PIS) [Participant Information Sheet]	2	08 June 2018
Summary CV for Chief Investigator (CI) [CV]		
Summary of any applicable exclusions to sponsor insurance (non-NHS sponsors only) [Public Liability Insurance Certificate]		05 September 2017
Validated questionnaire [Validated questionnaires]		

Statement of compliance

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

After ethical review

Reporting requirements

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and

the application procedure. If you wish to make your views known please use the feedback form available on the HRA website: http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

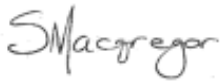
HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at http://www.hra.nhs.uk/hra-training/

18/WS/0114	**Please quote this number on all correspondence**

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

Yours sincerely



for
Dr Stewart Campbell
Chair

Enclosures: "After ethical review – guidance for researchers"

Copy to: Dr Thomas Webb, University of Sheffield
Ms Beverley Lowe, Lancashire Care NHS Foundation Trust

Appendix H

Evidence of Ethical Approval for StratCare trial by Health Research Authority

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

Dr Jaime Delgadillo
Lecturer in Clinical Psychology
Clinical Psychology Unit,
University of Sheffield
Cathedral Court, Floor F
1 Vicar Lane, Sheffield
S1 1HD

Email: hra.approval@nhs.net
Research-permissions@wales.nhs.uk

26 July 2018

Dear Dr Delgadillo

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Pragmatic randomised controlled trial of a stratified care model for depression and anxiety**
IRAS project ID:	**247945**
Protocol number:	**152958**
REC reference:	**18/LO/1116**
Sponsor	**Sheffield University**

I am pleased to confirm that HRA and Health and Care Research Wales (HCRW) Approval has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

How should I continue to work with participating NHS organisations in England and Wales?
You should now provide a copy of this letter to all participating NHS organisations in England and Wales, as well as any documentation that has been updated as a result of the assessment.

Following the arranging of capacity and capability, participating NHS organisations should **formally confirm** their capacity and capability to undertake the study. How this will be confirmed is detailed in the "*summary of assessment*" section towards the end of this letter.

You should provide, if you have not already done so, detailed instructions to each organisation as to how you will notify them that research activities may commence at site following their confirmation of capacity and capability (e.g. provision by you of a 'green light' email, formal notification following a site initiation visit, activities may commence immediately following confirmation by participating organisation, etc.).

It is important that you involve both the research management function (e.g. R&D office) supporting each organisation and the local research team (where there is one) in setting up your study. Contact details of the research management function for each organisation can be accessed here.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within the devolved administrations of Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) has been sent to the coordinating centre of each participating nation. You should work with the relevant national coordinating functions to ensure any nation specific checks are complete, and with each site so that they are able to give management permission for the study to begin.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?
The document "*After Ethical Review – guidance for sponsors and investigators*", issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

I am a participating NHS organisation in England or Wales. What should I do once I receive this letter?
You should work with the applicant and sponsor to complete any outstanding arrangements so you are able to confirm capacity and capability in line with the information provided in this letter.

The sponsor contact for this application is as follows:

Name: Dr Thomas Webb
Email: T.Webb@sheffield.ac.uk

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **247945**. Please quote this on all correspondence.

IRAS project ID	247945

Yours sincerely

Thomas Fairman
HRA Assessor

Email: hra.approval@nhs.net

Copy to: *Dr Thomas Webb, Sheffield University, (Sponsor Contact)*
Ms Beverley Lowe, Lancashire Care NHS Foundation Trust, (Lead NHS R&D Contact)

IRAS project ID	247945

List of Documents

The final document set assessed and approved by HRA and HCRW Approval is listed below.

Document	*Version*	*Date*
Covering letter on headed paper [Cover letter]		25 May 2018
Covering letter on headed paper [Cover letter]		08 June 2018
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [StratCare Insurance Certificate]		25 May 2018
HRA Schedule of Events [Schedule of events]	1.0	08 June 2018
HRA Statement of Activities [Statement of activities]	1.0	08 June 2018
Instructions for use of medical device [StratCare App: Instructions for use of non-medical device (Simply stated as MAY18)]	1	01 May 2018
IRAS Application Form [IRAS_Form_01062018]		01 June 2018
Letter from funder [Funding letter]		27 May 2018
Letter from sponsor [Sponsorship confirmation letter]		23 May 2018
Letter from statistician [Statistical review]		19 May 2018
Other [HRA definition of NON-MEDICAL device]		
Other [StratCare_Patients_Consent_process_v1]	1	04 July 2018
Other [StratCare_Patients_Info_Sheet_v1]	1	04 July 2018
Other [StratCare_Therapists_Consent_form_v3]	3	04 July 2018
Other [StratCare_Trial_Protocol_v2]	2	11 July 2018
Other [Cover letter 11.07.18]		11 July 2018
Participant information sheet (PIS) [Participant Information Sheet]	2	08 June 2018
Summary CV for Chief Investigator (CI) [CV]		
Summary of any applicable exclusions to sponsor insurance (non-NHS sponsors only) [Public Liability Insurance Certificate]		05 September 2017
Validated questionnaire [Validated questionnaires]		

IRAS project ID	247945

Summary of assessment

The following information provides assurance to you, the sponsor and the NHS in England and Wales that the study, as assessed for HRA and HCRW Approval, is compliant with relevant standards. It also provides information and clarification, where appropriate, to participating NHS organisations in England and Wales to assist in assessing, arranging and confirming capacity and capability.

Assessment criteria

Section	Assessment Criteria	Compliant with Standards	Comments
1.1	IRAS application completed correctly	Yes	The sponsor has confirmed that they do not consider that this is a study of a medical device requiring notification to the MHRA.
2.1	Participant information/consent documents and consent process	Yes	No comments
3.1	Protocol assessment	Yes	No comments
4.1	Allocation of responsibilities and rights are agreed and documented	Yes	The sponsor has submitted the HRA Statement of Activities and intends for this to form the agreement between the sponsor and study sites. The sponsor is not requesting, and does not require any additional contracts with study sites.
4.2	Insurance/indemnity arrangements assessed	Yes	Where applicable, independent contractors (e.g. General Practitioners) should ensure that the professional indemnity provided by their medical defence organisation covers the activities expected of them for this research study
4.3	Financial arrangements assessed	Yes	External study funding has been secured from MindLife UK Ltd. Study funding will be provided to sites, as detailed at Schedule 1 of the Statement of Activities.

IRAS project ID	247945

Section	Assessment Criteria	Compliant with Standards	Comments
5.1	Compliance with the Data Protection Act and data security issues assessed	Yes	No comments
5.2	CTIMPS – Arrangements for compliance with the Clinical Trials Regulations assessed	Not Applicable	No comments
5.3	Compliance with any applicable laws or regulations	Yes	No comments
6.1	NHS Research Ethics Committee favourable opinion received for applicable studies	Yes	No comments
6.2	CTIMPS – Clinical Trials Authorisation (CTA) letter received	Not Applicable	No comments
6.3	Devices – MHRA notice of no objection received	Not Applicable	No comments
6.4	Other regulatory approvals and authorisations received	Not Applicable	No comments

Participating NHS Organisations in England and Wales

This provides detail on the types of participating NHS organisations in the study and a statement as to whether the activities at all organisations are the same or different.
All participating NHS organisations will undertake the same study activities. There is therefore only one study site 'type' involved in the research. The Chief Investigator or sponsor should share relevant study documents with participating NHS organisations in England and Wales in order to put arrangements in place to deliver the study. The documents should be sent to both the local study team, where applicable, and the office providing the research management function at the participating organisation. Where applicable, the local LCRN contact should also be copied into this correspondence. If chief investigators, sponsors or principal investigators are asked to complete site level forms for participating NHS organisations in England and Wales which are not provided in IRAS or on the HRA or HCRW websites, the chief investigator, sponsor or principal investigator should notify the HRA immediately at hra.approval@nhs.net, or HCRW at Research-permissions@wales.nhs.uk. We will work with these organisations to achieve a consistent approach to information provision.

Principal Investigator Suitability

This confirms whether the sponsor position on whether a PI, LC or neither should be in place is correct for each type of participating NHS organisation in England and the minimum expectations for education, training and experience that PIs should meet (where applicable).
A Principal Investigator should be appointed at study sites. GCP training is not a generic training expectation, in line with the HRA/HCRW/MHRA statement on training expectations.

HR Good Practice Resource Pack Expectations

This confirms the HR Good Practice Resource Pack expectations for the study and the pre-engagement checks that should and should not be undertaken
As a non-commercial study undertaken by local staff, it is unlikely that letters of access or honorary research contracts will be applicable, except where local network staff employed by another Trust (or University) are involved (and then it is likely that arrangements are already in place). Where arrangements are not already in place, network staff (or similar) undertaking any of the research activities listed in A18 or A19 of the IRAS form (except for administration of questionnaires or surveys), would be expected to obtain an honorary research contract from one NHS organisation (if university employed), followed by Letters of Access for subsequent organisations. This would be on the basis of a Research Passport (if university employed) or an NHS to NHS confirmation of pre-engagement checks letter (if NHS employed). These should confirm enhanced DBS checks, including appropriate barred list checks, and occupational health clearance. For research team members only administering questionnaires or surveys, a Letter of Access based on standard DBS checks and occupational health clearance would be appropriate.

Other Information to Aid Study Set-up

This details any other information that may be helpful to sponsors and participating NHS organisations in England to aid study set-up.
The applicant has indicated that they do intend to apply for inclusion on the NIHR CRN Portfolio.

Appendix I

Sample Characteristics of Propensity Score Matched Sample

Table I1

Characteristics of Propensity Score Matched Sample

	Full Sample (n = 200)	CBT (n = 100)	CfD (n = 100)	Comparison of CBT and CfD sample Test statistics (d. f.)	p
Demographic/ Characteristic	Mean (SD)/%	Mean (SD)/%	Mean (SD)/%		
Demographics					
Age (years)	40.39 (14.2)	38.93 (13)	41.85 (15.2)	U = 4467	.193
Female	139 (69.5%)	68 (68%)	71 (71%)	χ^2 (1) = .212	.645
White	194 (97%)	91 (91%)	95 (95%)	χ^2 (1) = 1.23	.268
Unemployed	84 (42%)	28 (28%)	19 (19%)	χ^2 (1) =2.253	.133
Clinical characteristics					
Primary diagnosis					
Affective disorder	93 (46.5%)	24 (24%)	69 (69%)	χ^2 (1) = 40.699	< .001*
Anxiety disorder	53 (26.5%)	37 (37%)	16 (16%)	χ^2 (1) = 11.321	< .001*
Mixed (Affective and Anxiety disorder)	14 (7%)	7 (7%)	7 (7%)	χ^2 (1) = 0	1
Other/ Missing	31 (15.5%)	28 (28%)	8 (8%)	χ^2 (1) = 23.86	< .001*
Prescribed medication	113 (39.5%)	54 (54%)	59 (59%)	χ^2 (1) = .046	.829
Comorbid long-term medical illness	35 (17.5%)	20 (20%)	15 (15%)	χ^2 (1) = 1.073	.3
Disability	18 (9%)	10 (10%)	8 (8%)	χ^2 (1) = .293	.589

Table I1 (continued)

SAPAS score (initial assessment)	4.14 (1.5)	4.15 (1.5)	4.12 (1.5)	U = 4931	.863
Complex Cases	61 (30.5%)	34 (34%)	27 (27%)	χ^2 (1) = 1.156	.282
HIT session number	8 (5.2)	8.45 (5.9)	7.55 (4.3)	U = 4996	.993
HIT baseline PHQ-9	14.35 (5.9)	13.63 (6.5)	15.07 (5.1)	t(198) = 1.737	.084
HIT baseline PHQ-9 score ≥ 10	156 (78%)	71 (71%)	85 (85%)	χ^2 (1) = 5.711	.017*
HIT baseline GAD-7	13.09 (5.3)	12.86 (5.6)	13.31 (4.9)	t(198) = .602	.548
HIT baseline WSAS	20.61 (9.2)	19.12 (9.96)	22.1 (8.2)	t(198) = 2.304	.022*
Final PHQ-9	8.35 (6.2)	8.83 (6.9)	7.87 (5.4)	U = 4812	.645
Final GAD-7	7.74 (5.6)	8.04 (5.98)	7.38 (5.2)	U = 4803	.63
Final WSAS	14.02 (10.1)	13.29 (10.98)	14.74 (9.097)	U = 4387	.134

Note. SD = Standard deviation, SAPAS = Standardized Assessment of Personality – Abbreviated Scale, PHQ-9 = Patient Health Questionnaire-9, GAD-7 = Generalised Anxiety Disorder-7, WSAS = Work and Social Adjustment Scale. U = Mann-Whitney U test, t – Independent t-test, χ^2 = Chi-Square test of independence, * indicates statistically significant difference between CBT and CfD sample.

Appendix J

Reliable Improvement and Reliable Deterioration Rates for Propensity Score Matched Sample

Table J1

Overview of reliable improvement and deterioration of Propensity Score Matched Sample

	Total *n* (%)	CBT *n* (%)	CfD *n* (%)	Chi-Square Test of Independence
Standard and complex patients				
Sample size	200 (100%)	100 (100%)	100 (100%)	
Reliable Improvement	104 (52%)	42 (42%)	62 (62%)	χ^2 (1) = 8.013, p = 0.05
Reliable Deterioration	7 (3.5%)	1 (1%)	2 (2%)	χ^2(1) = 0.338, p = 0.561
Complex cases				
Sample Size	61 (100%)	34 (100%)	27 (100%)	
Reliable Improvement	34 (55.7%)	16 (47%)	18 (66.6%)	χ^2(1) = 2.345, p = 0.126
Reliable Deterioration	2 (3.3%)	1 (2.9%)	(3.7%)	-*

Note. Chi-square test of independence comparing CBT and CfD sample. *Assumptions for Chi-Square test not met.