Understanding the Role of Pre-Psychological Therapy Outcome Expectations, Therapy Preference and Wait Times on Non-attendance at the First Treatment Appointment and Clinical Outcomes

Amareen Hussain

A thesis submitted in partial fulfilment of the requirements for the degree

of Doctorate in Clinical Psychology

Clinical and Applied Psychology Unit - School of Psychology

The University of Sheffield

May 2025

Declaration

This thesis is submitted in partial fulfilment of the Doctorate in Clinical Psychology award at the University of Sheffield. I, the author, confirm this thesis is my own work, and all other sources have been referenced accordingly. I am aware of the University of Sheffield guidance on unfair means (https://www.sheffield.ac.uk/study-skills/assessment/academic-integrity). I declare this thesis has not been submitted for any other degree or to any other institution.

Structure and Word Count

Lay Summary: 298

Section One: Literature Review

Excluding references and tables: 7,391

Including references and tables: 11,045

Section Two: Empirical Study

Excluding references and tables: 7,652

Including references and tables: 11,108

Total word count

Excluding references and tables: 15,043

Including references and tables: 22,153

Lay Summary

Depression is a common mental health condition, affecting over 280 million individuals worldwide and can include symptoms such as low mood, hopelessness, and changes in appetite. For adults with depression, psychological therapy is one of the endorsed treatment approaches, with various therapies available such as cognitive behavioural therapy (CBT) and person-centred experiential therapy (PCET). However, not everyone benefits from receiving treatment, whilst many individuals drop-out of therapy. Identifying factors related to poorer outcomes and higher risk of non-attendance is crucial for patient-focused research. Potential contributing factors influencing engagement and therapeutic success could include outcome expectancy, treatment preference, and wait times.

The first chapter is a systematic review/meta-analysis that investigated the relationship between pre-treatment outcome expectancy and post-treatment outcomes in adults receiving CBT for depression. The existing literature was searched systemically, using various databases alongside backward and forward citation search, yielding 11 studies. Ten of the studies reported pre-treatment outcome expectancy to influence post-treatment outcomes, showing higher level of pre-treatment outcome expectancy to be associated with fewer symptoms of depression following therapy. Only one study showed a potential contrasting effect. More research is needed within the area in order to conduct a further comprehensive review.

The second chapter explored whether pre-treatment outcome expectancy, treatment preference, and wait times predicted non-attendance at the first treatment appointment and post-treatment clinical outcomes, in clients with depression, and

whether type of therapy received (PCET or CBT) was a moderating factor. A secondary analysis was conducted on data obtained from a completed trial within an NHS Talking Therapies service. High pre-treatment outcome expectancy was associated with better clinical outcomes but also associated with a higher risk of non-attendance at the first treatment appointment. Whether or not clients received their preferred therapy and how long they waited for therapy did not predict non-attendance nor clinical outcomes. Therapy modality was not a moderating factor.

Acknowledgements

I would like to begin by thanking God for blessings, guidance and strength throughout the journey.

I would like to express my gratitude to my research supervisors Professor Gillian Hardy and Professor Michael Barkham, for your guidance, reassurance, support and invaluable feedback throughout the process. Also, thank you to Dr. Dave Saxon, whilst you retired prior to the project being completed, the support provided in the early stages of the thesis, specifically around statistical analysis was a strong foundation for my work. An extended thank you to Dr. Ben Lorimer, for your continued statistical support and guidance, allowing me to successfully complete the project, following Dave's retirement.

Thank you to everyone who participated in the PRaCTICED trial - your involvement made this project possible.

I would like to thank my family and friends for their unwavering belief, motivation, encouragement and love throughout this journey. I am truly blessed to have you by my side. A special thank you to my grandma for her constant love and prayers. Your heartfelt prayers have been a source of strength and guidance.

And finally, a huge thank you to my DclinPsy cohort! Thank you for all the incredible support, love, shared experiences and unforgettable memories. I have learnt so much from every one of you. I am very lucky to have shared this experience with such an amazing group of people.

Table of Contents

Part One: Literature Review

Is Pre-Treatment Outcome Expectancy Associated with Post-Treatment Outcomes in Cognitive Behavioural Therapy for Depression in Adults? A Systematic Review and Meta-Analysis

Abstract

Objectives: Depression is a leading cause of disability and cognitive behavioural therapy (CBT) is a recommended therapy, though outcomes vary. One psychological factor linked to the success of therapy is outcome expectancy - the client's prognosis of how effective therapy will be for them. Whilst two previous reviews have examined the association between outcome expectancy and post-treatment outcomes, no review has specifically studied this in the context of CBT for depression. This systematic review aims to address this gap, by evaluating whether pre-treatment outcome expectancy is associated with post-treatment outcomes in CBT for depression.

Method: A systematic review (PROSPERO; CRD42024621516) was conducted, in line with the PRISMA guidelines. Systematic searches were conducted using Scopus, PsycINFO, and MEDLINE from inception to 2nd December 2024. Grey literature, in the form of dissertations and theses was searched using The Lens and ProQuest, from inception to 12th January 2025, alongside backward and forward citation search. Quantitative studies examining the association between pre-treatment outcome expectations and post-treatment outcome were included. Each included study underwent quality appraisal using the Critical Appraisal Skills Programme (CASP) checklists. Meta-analysis and narrative summary were used to summarise findings.

Results: Eleven studies were included in this review. Seven were included in the meta-analysis which revealed a small to moderate negative association between pre-treatment outcome expectations and post- treatment outcomes (r = - 0.21, 95% CI [-0.32, - 0.08]. This indicated that individuals who had higher expectations for

treatment report lower depressive symptoms following treatment. Four studies were narratively summarised, which further supported this trend, with three demonstrating a similar negative association, whereas one study reported a small positive relationship.

Conclusion: These findings show the important role pre-treatment outcome expectancy can play in outcomes in clients receiving CBT for depression. Future research should continue to assess the impact of outcome expectations on CBT outcomes for depression to expand on the limited existing evidence-base. Further research should study the association between pre-treatment outcome expectancy and follow-up outcomes.

Practitioner Key Points:

- Pre-treatment outcome expectancy is associated with better outcome in adults receiving CBT for depression.
- Outcome expectancy should be assessed prior to therapy commencing, through a brief outcome expectancy measure or through verbal discussion.
- Strategies should be implemented to address outcome expectations of treatment, to enhance patients' belief in therapy and to support better treatment outcomes.

Key Words: Depression, CBT, pre-treatment outcome expectancy, post-treatment outcomes

Introduction

Depression, a common mental health disorder is a global cause of disability impacting over 300 million people globally. The World Health Organisation recognises depression as the primary cause of disability, accounting for 4.3% of the global disease burden and estimates that by 2030, it will be the leading contributor (World Health Organisation, 2023). This common mental health disorder is marked by symptoms of persistent sadness or irritability, reduced interest and a range of physical and emotional symptoms (Tylee & Gandhi, 2005). It alters an individual's thoughts, feelings and behaviours, impairing daily functioning (Gotib & Joormann, 2010), and is associated with an increased risk of suicide (Turecki & Brent, 2017). Depression often co-occurs alongside other physical and mental health issues, increasing the risk of developing chronic and acute physical health problems (Mezuk et al., 2008; Pan et al., 2012; Win et al., 2011) and substance use disorders (Conner et al., 2009).

Depression is frequently treated using cognitive behavioural therapy (CBT). CBT is the recommended first-line intervention for depression by the National Institute for Health and Care Excellence (NICE, 2022). CBT has established itself as the cornerstone treatment for depression and has demonstrated efficacy in alleviating depressive symptoms (Hofmann et al., 2012; Werson et al., 2022). Whilst it has demonstrated its effectiveness in the treatment of depression, it should be noted that 15%-45% of clients do not experience significant improvement (Hansen et al., 2002), whilst up to 10% of client in fact deteriorate (Lambert & Ogles, 2004). Several factors contribute to this variability, including patient, therapist, and treatment factors (Branson et al., 2015; Hamilton & Dobson, 2002).

One important but often overlooked factor influencing therapy outcomes is patient expectations. Patient expectations have been identified to be a vital component of successful therapy outcome. Constantino et al. (2021) and Frank (1961) outline that therapy works the most effectively when clients have hope that the situation will improve. Psychotherapy expectations can be categorised into several distinct types, and there is often confusion differentiating between them (Delsignore & Schnyder, 2007). Outcome expectations, the focus of this review, refer to a patient's prediction about how they will respond to treatment. In other words, it is their belief about how helpful a treatment will be in addressing their issues (Constantino et al., 2018)

Historically there has been debate relating to whether treatment credibility and outcome expectancy are the same construct, as outcome expectancy may partially develop based on how credible a treatment seems (Hardy et al., 1995). However, recently clear distinctions have been made, now recognising them as separate constructs. Treatment credibility refers to the extent to which treatment seems logical and makes sense to the client (Newman & Fisher, 2010). Whereas outcome expectancy is the belief how well the treatment will work for them personally (Constantino et al., 2018). Outcome expectancy can form prior to one having knowledge about the treatment or any direct interaction with the practitioner, whereas credibility is typically shaped by experience and observation (Schulte, 2008). Furthermore, credibility has been thought to come from logical thought process, whilst outcome expectancy is known to come from feelings or emotions (affective processes; Devilly & Borkovec, 2000).

In line with identifying predictors of response to therapy, pre-treatment outcome expectancy has been found to be associated with improved outcomes. A

meta-analysis by Constantino et al. (2011) of 46 independent samples found a significant association between higher pre-treatment or early treatment outcome expectancies and post-treatment outcomes (weighted r = .12 or d = .24). An updated meta-analysis by Constantino et al. (2018) consisting of 81 independent samples revealed an association of r = .18 or d = .36. These findings are underpinned by expectancy theory, positing that people are motivated to work towards a goal if they believe their effort will result in good outcome, and a valued reward. This further shows that emotional and cognitive processes linked to positive outcome expectations play an important role in therapeutic success (Vroom, 1964).

Whilst the previous reviews contribute to understanding the relationship between outcome expectancy and outline link, they did not focus exclusively on CBT for depression. Instead, they focused on a broad range of therapies and disorders. Whilst Constantino et al. (2018) aimed to address the limitations of the Constantino et al. (2011) review, issues remained, including the failure to appraise the quality of studies. This review aims to address this gap by critically appraising the studies included, to provide a more reliable understanding of the impact of pre-treatment outcome expectancy on post-treatment outcomes. Constantino et al. (2018) attempted to distinguish outcome expectancy from treatment credibility. However, their inclusion criteria allowed studies referencing outcome expectancy as their independent variable, even when calculated in a manner that conflated the two constructs, such as utilising the Credibility/Expectancy Questionnaire (CEQ) total score. This is problematic, as it makes it challenging to draw clear conclusions about the specific role of outcome expectancy. This review adopts a more targeted approach. It only includes studies that define and measure outcome expectancy as a separate construct, such as those using the expectancy subscale of CEQ or similar

tools. Constantino et al. (2018) did not find treatment modality or diagnosis to be a significant moderator, however within their review they do not report statistical data for these non-significant results. This lack of clarity makes it difficult to evaluate the robustness of their conclusions.

Additionally, the two previous reviews excluded grey literature, which can provide more comprehensive coverage and reduce biases. The current review addresses this limitation by including such sources for a more comprehensive analysis. Lastly, this review is more up to date, including studies published after 2017. To the author's knowledge, this is the first review to explicitly study the association between pre-treatment outcome expectancy and post-treatment outcome, specifically for CBT, the recommended therapy for depression.

Aim

The aim of the review is to investigate if client pre-treatment outcome expectancy is associated with post-treatment outcomes in CBT for depression. Focusing on CBT, this review aims to enhance the understanding of how pre-treatment outcome expectancy beliefs about treatment effectiveness, can shape the success of this widely used and well-established evidence-based treatment for depression.

Clinical Implications

Research exploring factors that contribute to the effectiveness of CBT for depression, highlights the importance of considering psychological factors beyond treatment protocols and techniques. This is central to patient-focused research (Lutz, 2002). Systematic reviews are considered vital and widely accepted in research, as they provide a rigorous and reproducible method for synthesising evidence from

various studies. Systematic reviews decrease bias by consolidating findings from different studies, making findings more trustworthy, and therefore they offer a comprehensive understanding of a particular field. Systematic reviews can inform future research, help clinicians stay current with the latest evidence and inform clinical practice.

If outcome expectancy is found to be associated with treatment success, then clinicians may consider exploring these expectations early on in therapy. Dependent on the nature of the relationship between expectancy and outcome, this may include utilising strategies to address expectations.

Method

The review has been completed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). The guidelines outlined by PRISMA can be found in Appendix A.

Protocol Registration

The protocol for this systematic review was registered on the international prospective register of systematic reviews (PROSPERO; Reference: CRD42024621516). This is available at https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=621516

Search Strategy

Initially, a brief hand search using Google Scholar was conducted to ensure the feasibility of the proposed review. This further allowed the author to identify key words appearing within the relevant articles. Where applicable, the reviews by Constantino et al. (2011; 2018) were used to inform the search strategy. Based on

the initial searches and eligibility criteria, the author developed the search syntax. The search terms were evaluated for relevance by two research experts within the field (the author's supervisors) and a liaison librarian for science, all affiliated with the University of Sheffield, England. Following discussion and mutual agreement, the search terms were revised and finalised. The search syntax can be found in Table 1.

The search terms were kept broad to ensure no relevant papers were missed at this stage. Broad terms related to 'cognitive behavioural therapy, and outcome expectancy' were combined. It should be noted, while the review is focused on depression, including 'depression' as a search term further restricted the number of results. To ensure no appropriate studies were missed, a broader search strategy was employed, allowing a comprehensive screening process. Although, this review focuses on outcome expectancy, and not 'treatment expectancy' the latter was still included within the search strategy, due to inconsistency in terminology. Historically, these terms have been used interchangeably or incorrectly, despite their distinct meanings. Therefore, a broader search approach was taken, with the screening process allowing the identification and selection of appropriate studies. Wildcards were used to capture variations in word endings. The terms were further combined using proximity operators and Boolean operators ("OR and "AND"). Boolean operators refine the systematic database search, combining 'OR' to account for variations in individual construct terms with 'AND' to narrow the focus of the results.

A systematic literature search was conducted across five electronic databases (SCOPUS, PsycINFO, MEDLINE, The Lens and ProQuest). Scopus, PsycINFO, and MEDLINE were searched for relevant studies from inception to 2nd December 2024. The Lens and ProQuest facilitated the searching of grey literature in the form of dissertations and theses, which will have not been peer-reviewed nor formally

published, from inception to 12th January 2025. All grey literature documents within ProQuest and Lenswas screened. Both The Lens and ProQuest were selected to ensure a thorough search of grey literature, due to differences in their respective repositories. The Lens provides access to open-access grey literature, primarily from publicly available sources. Whilst ProQuest offers a broader collection of theses and dissertations, including those not covered by The Lens.

Table 1

Search Syntaxes

Database	*Search Terms*
Scopus	(TITLE-ABS-KEY) "outcome expect*" *OR* "treatment expect*" *OR* (outcome W/3 expect*) *OR* (treatment W/3 expect*) *OR* (expect* W/3 symptom W/3 improvement) *OR* (expect* W/3 symptom W/3 change) *AND* (TITLE-ABS-KEY) CBT *OR* "cognitive behavio* therap*" *OR* "cognitive therap*" *OR* "behavio* therap*" *AND* LIMIT TO LANGUAGE, “English"
PsycINFO	(ALL FIELDS) exp Cognitive Behavior Therapy/ *OR* "cognitive behavio* therap*" *OR* CBT *OR* "cognitive therap*" *OR* “behavio* therap* AND "outcome expect*" *OR* "treatment expect*" *OR* (outcome adj3 expect*) *OR* (treatment adj3 expect*) *OR* (expect* adj3 symptom adj3 change) OR (expect* adj3 symptom adj3 improvement) AND LIMIT TO LANGUAGE, “English"
Medline	(ALL FIELDS) exp Cognitive Behavior Therapy/ *OR* "cognitive behavio* therap*" *OR* CBT *OR* "cognitive therap*" *OR* “behavio* therap* AND "outcome expect*" *OR* "treatment expect*" *OR* (outcome adj3 expect*) *OR* (treatment adj3 expect*) *OR* (expect*

adj3 symptom adj3 change) *OR* (expect* adj3 symptom adj3 improvement) AND LIMIT TO LANGUAGE, “English"

The Lens: (TITLE - ABS - KEY OR FIELD OF STUDY) "Outcome expect*" OR "treatment expect*" OR "outcome expect"~3 OR "treatment expect*" ~3 OR "expect* symptom improvement"~3 OR "expect* symptom change"~3 AND CBT OR "cognitive behavio* therap*" OR "cognitive therap*" OR "behavio* therap*"

ProQuest: (TITLE-ABS-KEY): "Outcome expect*" OR "treatment expect*" OR outcome NEAR/3 expect* OR treatment NEAR/3 expect* OR expect* NEAR/3 symptom NEAR/3 improvement OR expect* NEAR/3 symptom NEAR/3 change) AND CBT OR "cognitive behavio* therap*" OR "cognitive therap*" OR "behavio* therap*"

Note: Boolean operators were employed to refine the systematic database search, combining ‘OR’ to account for variations in individual construct terms with ‘AND’ to narrow the focus of the results. MEDLINE uses *exp* for MeSH terms (subject-specific indexing), and PsycINFO uses *exp* for thesaurus-based indexing.

Eligibility Criteria

Applying the Population, Intervention, Comparison and Study (PICOS) framework (Eriksen & Frandsen, 2018), studies were considered eligible for inclusion if they fulfilled the criteria outlined in Table 2.

Table 2

Inclusion and Exclusion Criteria for Study Eligibility

	Inclusion Criteria	**Exclusion Criteria**
Population	Clients aged ≥ 18 years displaying depressive symptoms that meet the diagnostic criteria for depression (ICD-10 or DSM-V criteria), or diagnosis of depression identified via scoring above clinical threshold on a validated screening measure. Depression is the primary condition. Comorbid physical or mental health conditions permitted if depression is the primary condition and focus of treatment. Can be taking anti-depressant medication.	Clients aged < 18 years whose depressive symptoms do not meet the diagnostic criteria for depression or do not score above the clinical threshold on a validated screening measure. Depression is not the primary condition.
Intervention	CBT for depression (CBT, cognitive or behavioural therapy). CBT permitted, with or without pharmacological treatment. Standard CBT therapy or internet-based/self-help CBT therapy with meaningful contact with therapist. Therapy delivered face-to-face or remotely. Can be 1-1 therapy or group therapy intending to offer least ≥ 2 intervention sessions	Treatment that does not align with the principles of cognitive and behavioural therapy and treatment is not targeting depression. Internet based/self-help therapy without meaningful contact with therapist.
Comparator	Not applicable	.

Outcome	Validated measures of depression administered at baseline and post-intervention (can be self-reported or clinician administered)	Depression symptoms not measured using validated measure.
	A self-reported quantitative measure of outcome expectancy administered pre-treatment with clear distinction from treatment credibility.	Outcome expectancy measured not administered pre-treatment or is not self-reported Outcome expectancy measured retrospectively or if outcome expectancy was manipulated.
		Studies where outcome expectancy and credibility or other constructs are conflated or treated synonymously.
	Quantitative association measured between pre-treatment outcome expectancy and post-treatment outcome.	Outcome expectancy measure is too broad (focuses on self-efficacy, hopelessness etc.)
		No statistical analysis for association between pre-treatment outcome expectancy and post-treatment outcome.
Setting	Any setting (inpatient/outpatient)	
Study Design	Quantitative research - no restrictions based on study design other than not being case studies.	Qualitative studies - reviews - book chapters - conference proceedings - presentations - media articles - case studies
	Published literature and grey literature in the form of dissertations and theses	
	Studies published in English Language	Studies published in non-English language

Note. Diagnostic and Statistical Manual of Mental Disorders 5th Edition (DSM-V), International Classification of Diseases 10th revision (ICD-10).

Study Selection

The articles identified through the data search were exported into EndNote - Version 21, a reference management software program developed by Clarivate (2013). EndNote is commonly used in systematic reviews for organising, managing and deduplicating references. Initially, Scopus, PsycINFO, MEDLINE database results were exported and merged, and any duplicates were removed. Following the removal of duplicates, the titles and abstracts of the remaining studies were reviewed independently by the primary reviewer, against the inclusion and exclusion criteria. To minimise the risk of inappropriate exclusion, for articles where uncertainty existed, such as those discussing treatments for depression or where the treatment approach was not explicitly specified, they were included in the full-text review. Additionally, studies where it was unclear whether pre-outcome expectancy was considered in relation to its association with treatment outcome were also included in the full-text review. The Lens and ProQuest database searches for grey literature were exported separately at a later data, and the same screening process was applied.

Following the title and abstract screening, all full texts of the remaining articles, were sourced again and assessed against the eligibility criteria by the primary reviewer. Subsequently, the primary reviewer conducted forward and backward citation searching for all included studies to identify any further articles that may have been missed during the database search. To ensure reliability and rigor, a second reviewer also independently double-checked and screened 64.1% of the papers that progressed to full-text screening (100% agreement).

Quality Assessment

The quality of eligible studies was assessed, using the Critical Appraisal Skills Programme (CASP) Checklists (CASP, 2022). CASP provides tools to assess the quality and rigor of a variety of study designs, including randomised control trials (RCTs), and cohort studies. The CASP checklist was adapted (omitted/added/re-phrased items), to ensure the quality appraisal was appropriate to the current systematic review. Appendix B and C illustrates the original CASP checklists which were used as templates, to create the adapted CASP checklists. In summary, for the cohort studies, the CASP checklist was adapted by removing questions relating to follow-up and including questions relating to the accuracy of outcome expectancy and depression measures. The CASP RCT checklist was also adapted. Changes included, adding questions relating to the accuracy of outcome expectancy and depression measures. Also, questions relating to randomisation of participants, whether participants and assessors were blind to intervention were omitted. Original CASP checklists alongside adapted checklists can be found in Appendix B, C and D.

The studies were evaluated against the appropriate checklist by the lead author. Ratings on the CASP are either ‘yes’ ‘no’ or cannot tell. CASP does not provide a total score. The number of ‘yes’ responses were totalled to facilitate the evaluation of study quality. A ‘yes’ response indicated higher quality, thus the greater the number of ‘yes’ responses, the higher the quality of the study. A scoring rubric was created to assess studies based on the CASP criteria, assigning one point for each met criterion and zero points for unclear or unmet criteria. For RCTs, this resulted in a quality score ranging from 0-9, which translated into 0-4: low quality, 5-7: moderate quality and 8-9: high quality ratings. For cohort studies, this resulted in quality score ranging from 0-12, which translated into 0-5, low quality, 6-9 moderate

quality and 10-12 as high quality. Appendix D outlines adapted checklists and the final quality appraisal ratings for all studies. All CASP criteria were assigned equal importance. 45.45% of the included studies were further selected at random and appraised independently by a secondary reviewer (the same reviewer who assessed a proportion of articles at full-text screening). Assessment of inter-rater reliability was conducted using Cohen's Kappa (Cohen, 1960) and this was categorised as moderate agreement, k = .60 (Landis & Koch, 1977). Given that values between .41 and .60 are categorised as moderate, .60 is on the higher end of this range, implying a relatively strong level of consistency between the two raters. All studies were included within the review to prevent selection bias and to provide a comprehensive overview of the evidence (Popay et al., 2006).

Data Extraction Procedure

For each publication, the primary researcher extracted and summarised the following information. Author (s), year of publication, country, study design, sample characteristics, details of the CBT modality, outcome measures of pre-treatment outcome expectancy and depression, appropriate findings and quality rating.

Data Analysis

Data Synthesis

The analysis entailed both a meta-analysis for studies with appropriate data available and narrative summary for those studies where the necessary meta-analytical data was not available.

Meta-analysis Strategy

A random effects meta-analysis was conducted using the Meta-Essential software (version 1.5; Surrmond et al., 2017), because high levels of variability within the meta-analysis was expected. A random effects meta-analysis assumes variability in the data and assumes that the true effect could differ across studies due to the heterogeneity between them. Thus, a random effects meta-analysis utilises the inverse-variance approach, adjusting study weights based on the degree of variation and heterogeneity between studies (Dettori et al., 2022). This decreases the likelihood of type one error occurring (Hunter & Schmidt, 2000). The meta-analysis was focused on the association between pre-treatment outcome expectancy and post-treatment depression outcomes. Pearson's *r* correlations were extracted from studies. When *r* was not reported, but other appropriate statistical information (for instance, means, standard deviations, odds ratios) was available, these were then used to convert to the same common metric of *r.* This was completed using the online Psychometrica calculator (Lenhard & Lenhard, 2002). Where appropriate statistics were not provided, the authors were contacted by email and given 10 days to respond. Once all effect sizes had been converted into the common metric to make them comparable, Meta-essentials converted these to Fisher's z to stabilise variances and to calculate a pooled effect size. Results were then converted back to Pearson's r for easier interpretation.

A forest plot was generated to visually represent the effect sizes and confidence intervals for each study, as well as the pooled effect size. The following guidelines for interpreting effect sizes were used: $r = .10$ for small effect, $r = .30$ for a moderate effect and $r = .50$ for a large effect (Cohen, 1988).

Heterogeneity assessing the variance between studies was examined using the Q statistic and the I^2 statistic. The Q statistic assesses the degree of variability among the pooled effect sizes. A statistically significant result indicates true heterogeneity across the effect sizes instead of variability that is due to pure error (Card, 2012). The I^2 statistic quantifies the proportion of total variability due to between-study heterogeneity, that is not due to sampling error. The I^2 values indicating differing levels of heterogeneity are as follow: 25% is seen as low, 50% is seen as moderate and 75% is considered high (Higgins et al., 2002). When a high level of heterogeneity is presented a moderator analysis may be warranted. Due to a limited number of studies (N=7) within the meta-analysis, it was decided neither a sub-group analysis nor a meta-regression would be conducted. It Is recommended a moderator analysis (sub-group or meta-regression) should only be performed when the meta-analysis consists of more than ten studies and when sub-groups consist of three or more studies (Card et al., 2012). Therefore, conducting moderator analyses was not deemed appropriate.

To further assess the robustness of the meta-analysis findings, a publication bias is recommended. Publication bias refers to the tendency for studies with larger effects to be more likely to be published, resulting in an upward bias in the summary effects. To assess and remove publication bias, Card (2012) recommends various methods, such as the inspection of funnel plots, the trim and fill method (Duval & Tweedie, 2000), fail-safe N (Rosenthal, 1979) and Egger's regression test (Egger et al.,1997)

Upon inspection of the funnel plot, publication bias is suggested when there looks to be an asymmetrical distribution (deviation from the expected triangular configuration) of the individual effect sizes within the meta-analysis (Card 2012;

Peters et al., 2008). In addition to visually inspecting the funnel plot, Egger's regression test (Egger et al.,1997) statistically assesses the asymmetry in the funnel plots. A significant value on the intercept test result would imply a risk of publication bias.

The trim-and-fill method (Duval & Tweedie, 2000) identifies and corrects for funnel plot asymmetry due to publication bias. The approach involves 'trimming' (removing) studies contributing to asymmetry. The trimmed funnel plot is then used to approximate the true centre of the funnel and impute hypothetically the missing studies (filling). An adjusted-bias corrected summary effect is then derived. Results are deemed to imply publication bias if they are not comparable to the initial values (Card, 2012).

Finally, to assess for publication bias, the Rosenthal (1979) fail safe N was calculated. This estimates the number of studies with non-significant results that would be required in the meta-analysis to overturn the overall significant effect to non-significant ($p < .05$). Rosenthal's (1979) guidelines were applied to interpret the fail-safe N, with an adequate value being one that exceeds $5k + 10$ (k = number of studies included in the review).

Results

Search Results

The study selection is illustrated in Figure 1 using a PRISMA diagram (Page et al., 2021). Firstly, a systematic search was conducted using three electronic databases: PsycINFO, Scopus, and Medline, yielding a total of 1,727 records. Upon removing duplicates (N=714) 1013 records remained. These records were screened by title and abstract according to the inclusion and exclusion criteria. For the

remaining articles (N=38), full text articles were retrieved and screened against the inclusion criteria. Following full-text screening twenty-eight studies were excluded. Reasons for excluding these studies are stated in Appendix E.

A further search through backward and forward citation search yielded three potentially relevant studies, one of which met the inclusion criteria and was further assessed at full-text screening and included within the review. A grey literature search was additionally conducted using Lens and ProQuest, but no additional studies were included from this. Combined, eleven studies were included for the systematic review and seven of these were included in the meta-analysis.

Study Characteristics

The characteristics of the included studies are outlined in Table 3. Of the 11 studies, five were conducted in the United States of America (USA), four in Europe, one in Australia and one in Canada. Five studies utilised a primary cohort design. Four studies were secondary analyses of original randomised controlled trials (RCTs), one study was a secondary analysis of a cohort study, one was a primary RCT.

The sample sizes ranged from 33 to 420 (M=217.55; SD = 182.39). Most samples recruited middle-aged adults (mean age of participants ranged from 34-45 years, with an overall range of 18 to 70 years of age). Most studies comprised a higher percentage of female participants. It should be noted for some studies only characteristics for the full sample were reported and not for the depression specific sub-sample. Notably, a few studies did not report specific study characteristics.

As stated in the inclusion criteria, all studies utilised CBT for depression, involving meaningful contact with the therapist. Therapy modalities included: group

CBT with combination of individual CBT (N=4), individual CBT (N=3), group CBT (N=2), with some studies incorporating internet-based CBT with therapist guidance (N=2).

Depression was measured using a variety of tools; four studies utilised self-reported Beck's Depression Inventory (BDI-II; Beck et al., 1996;), including one that used the BDI-II German version (Hautzinger et al., 1994). Three studies utilised self-reported Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001), and three studies measured depression using self-reported Centre for Epidemiological Studies Depression Scale-10 (CESD-10; Andresen et al., 1994). Finally, one study used the clinician rated Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960).

For expectancy, six studies used the Credibility and Expectancy questionnaire (CEQ; Devilly & Borkovec, 2000) and in all cases the expectancy component was extracted separately. Three studies opted to use a single-item expectancy measure, whilst one study opted to use the Outcome Expectancy Scale (OES; Ogrodniczuk & Sochting, 2010). One study used a two-item questionnaire based on the work of Borkovec and Nau (1972) and Devilly and Borkovec (2000).

Figure 1

PRISMA (2009) flow diagram. Adapted from Moher et al. (2009).

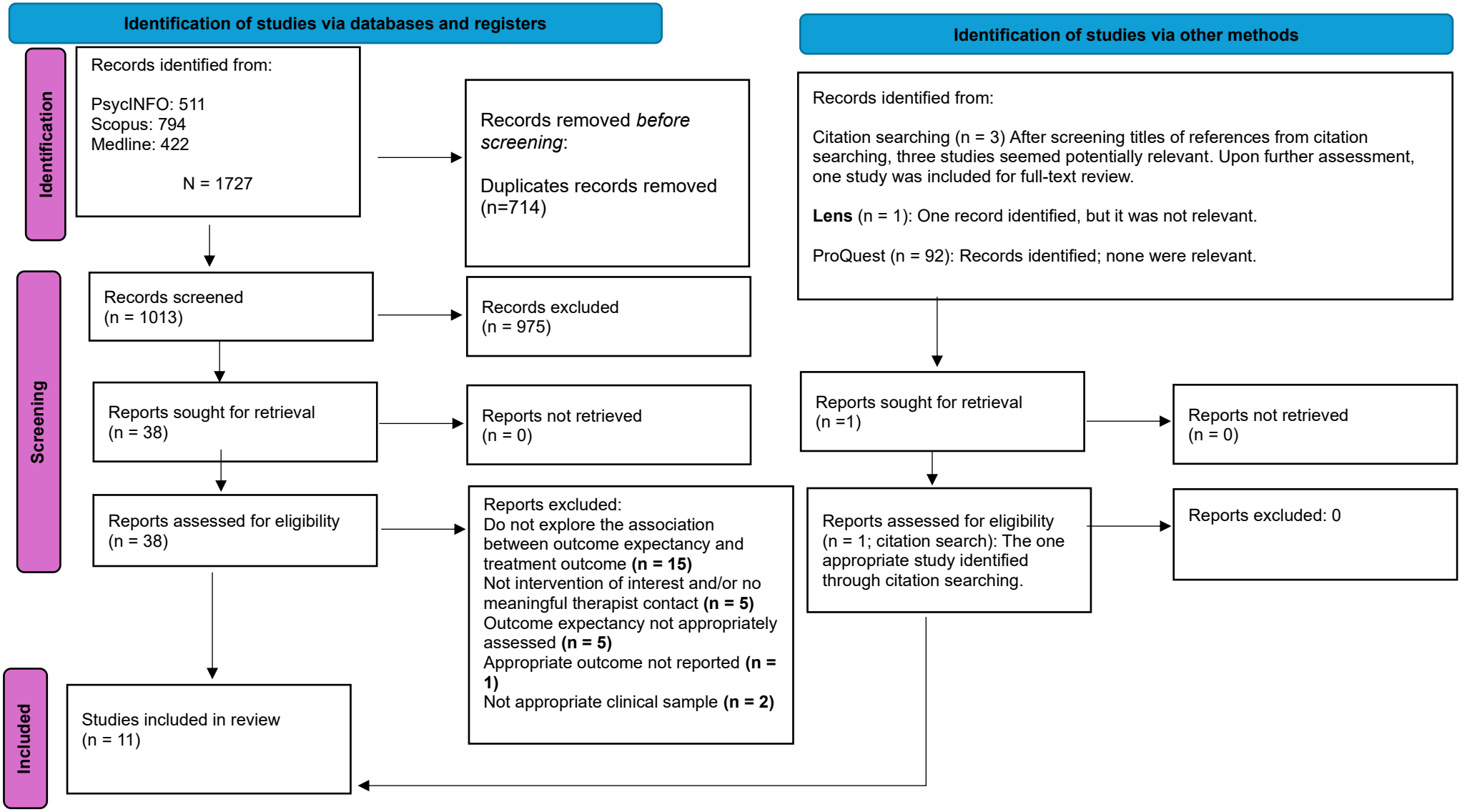


Table 3

Characteristics of Included Studies

Author (Year)	Country	Sample Size	Study Design	Mean Age (SD) [range]	Sex % Female	Therapy Modality	OE Measure	Depression Measure	Quality Rating Score
Beard et al. (2016)*	United States	Sub-sample of interest: N = 539 (Total study sample: N = 956)	Cohort	34.71	57%	Group CBT with 1-1 sessions	CEQ	CESD-10	High
Friedl et al. (2020)	France, Netherlands, Switzerland, Denmark	251	Secondary analysis of an original RCT study	41 (13.7)	68.2%	Individual CBT with internet-based CBT TAU: Standard CBT	CEQ	PHQ-9	Moderate
Høifødt et al. (2015)*	Norway	82	Secondary analysis of an original RCT study	36.0 (11.7) [18-65]	73%	Web- based CBT with 1-1 therapist guidance	Single item measure	BDI-II	Moderate
Ramondo et al. (2024)*	Australia	Sub-sample of interest: N=33 (Total study sample N = 66)	RCT	44.6 (12.27) [18-65]	76%	Group CBT	CEQ	PHQ-9	Moderate
Schindler et al. (2013)*	Germany	193	Cohort	38.6 (13.0)	68.4%	Individual CBT	Single item measure	BDI-II-German version	Low

Snippe et al. (2015)	Netherlands	Sub-Sample of interest N= 45 Total sample N= 91	Secondary analysis of an original RCT study	53.1 (11.8) [18-70]	49%	Individual CBT	2-item questionnaire (Borkovec & Nau, 1972; Devilly & Borkovec, 2000).	BDI-II	High
Vittengl et al. (2019)*	United States	152	Secondary analysis of an original RCT study	45 (13)	67%	1-1 CT	Single item measure	Clinician rated HRSD	Moderate
Visla et al. (2018)	Canada	91	Secondary analysis of an original cohort study	NS	NS	Group CBT	OES	BDI-II	Moderate
Webb et al. (2013)*	United States	Sub-sample of interest N= 420; Total sample: N=576	Cohort	NS	NS	Group CBT with 1-1 sessions	CEQ	CESD-10	High
Webb et al. (2014)*	United States	103	Cohort	36.02 (13.71) [18-68]	64%	Group CBT with 1-1 sessions	CEQ	CESD-10	Moderate
Webb et al. (2020)	United States	484	Cohort	34.0 [18-72]	59.9%	Group CBT with 1-1 sessions	CEQ	PHQ-9	Moderate

*Included in meta-analysis - RCT - Randomised controlled trial; CBT - Cognitive Behavioural Therapy; CT - Cognitive Therapy; CEQ - Credibility Expectancy Questionnaire; OES - Outcome Expectancy Scale; CESD-10 - Centre for Epidemiological Studies Depression Scale-10; PHQ-9 - Patient Health Questionnaire-9; BDI - Beck Depression Inventory; HRSD - Hamilton Rating Scale for Depression - NS; Not stated. Beard et al. (2016); Sample characteristics only available for full sample; Snippe et al. (2015); Sample characteristics of full sample available only (extracted from original study).

Methodological Quality

Ten studies were subjected to critical appraisal using to CASP cohort checklist. Of the 10, six were rated as being moderate quality, three as high quality, and one as weak quality. One study was appraised by using the CASP RCT checklist and was rated as having moderate quality. The quality scores for these studies ranged from five to eleven.

A key strength across all studies was the clear focused issue being stated, along with appropriate recruitment strategies, with a clear inclusion/exclusion criterion being reported. Depression outcomes were measured adequately with well-established measures. Most studies used well-established multi-item pre-treatment outcome expectancy measures, although three of those that used well-established expectancy measures, failed to report the psychometric properties of these within the paper. Three studies used single item outcome expectancy measures, and no psychometric properties of these were reported.

An area of weakness across most studies was the unclear reporting of whether important confounders were accounted for, while others did not account for these. For instance, some studies, were not clear or failed to account for confounding variables such as demographic, personal characteristics and technological familiarity. Moreover, several studies failed to report precision estimates, limiting the reliability of findings, making it difficult to draw firm conclusions. Where appropriate, some studies offered detailed clinical implications, whilst others provided vague or none, reducing the utility of the studies for informing clinical practice.

Meta-analysis

Seven studies were included in the meta-analysis to assess for the associations between pre-treatment outcome expectancy and post-treatment outcome. These seven studies were included in a random-effects model meta-analysis to examine pooled correlation coefficient. From the seven studies, seven individual effect sizes were obtained. Overall, the pooled sample size was 1522 participants.

Table 4 outlines the correlations, study coding, and results from the meta-analysis. A forest plot was generated (see Figure 2) which displays the effect size of each study, with their corresponding confidence intervals, the pooled correlation coefficient, and prediction intervals.

The meta-analysis revealed a small to moderate negative association between pre-treatment outcome expectancy and post-treatment depression outcomes (r = -.21, 95% CI [-.32, -.08]; z = -3.97, p < .001). This negative association suggests that higher pre-treatment outcome expectations are linked to lower levels of depression severity post-treatment. The prediction interval ranged from r = -.48 to r =.11. The tests assessing for heterogeneity of the effect sizes were significant, Q= 22.36, p = .001; I^2 = 73.17%, indicating moderate to high heterogeneity. Given the limited number of studies, moderator analyses were not conducted due to concerns about statistical power.

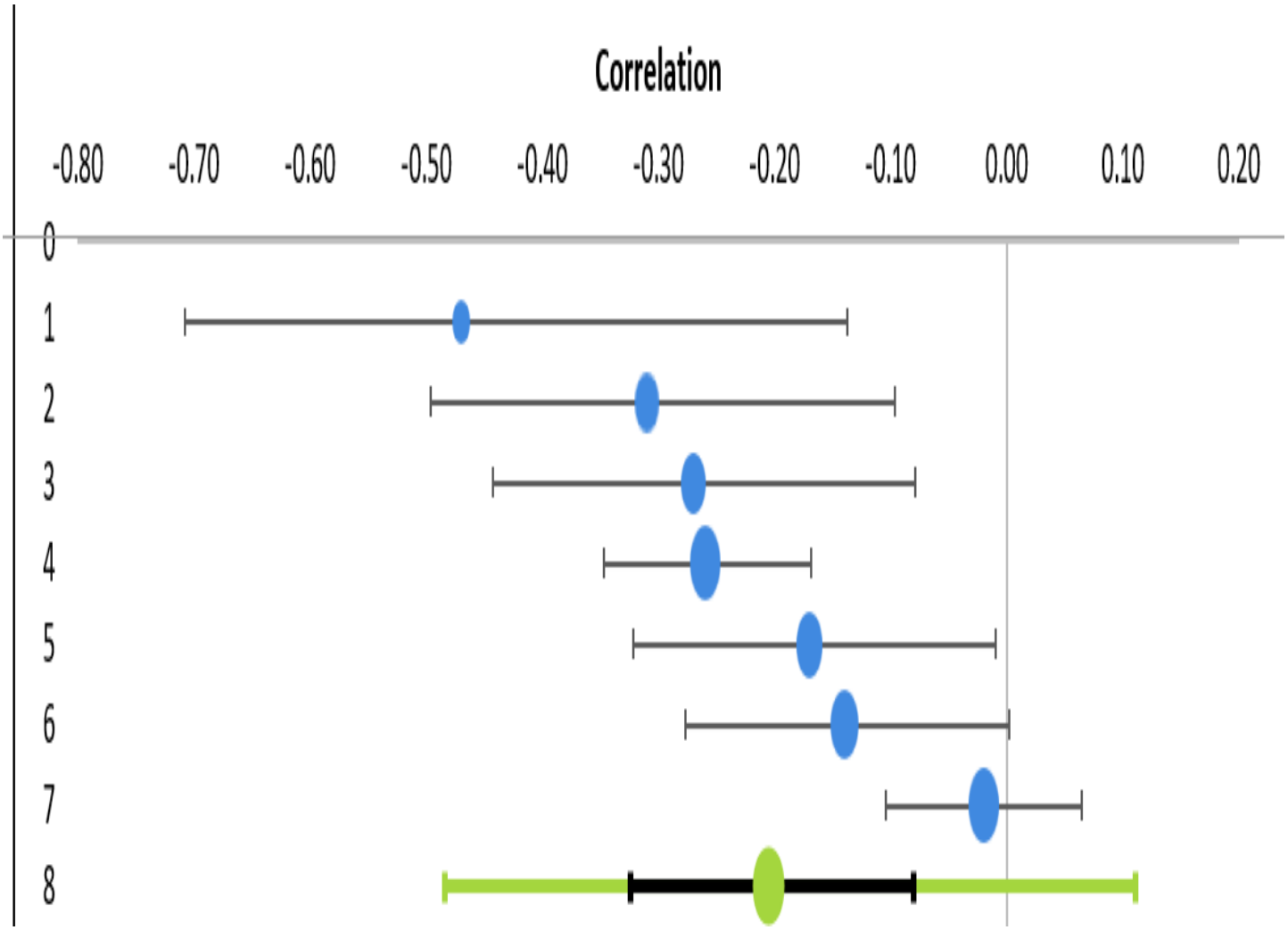
Table 4

Meta-Analysis Correlations

No.	Study	Correlation	95% CI		Weight
			Lower Limit	Upper Limit	
1	Ramondo et al. (2024)	-.47	-.71	-.14	6.57%
2	Høifødt et al. (2015)	-.31	-.50	-.10	11.60%
3	Webb et al. (2014)	-.27	-.44	-.08	12.87%
4	Webb et al. (2013)	-.26	-.35	-.17	18.73%
5	Vittengl et al. (2019)	-.17	-.32	-.01	14.89%
6	Schindler et al. (2013)	-.14	-.28	.00	15.99%
7	Beard et al. (2016)	-.02	-.10	.06	19.35%
8	**Overall**	**-.21**	**-.32**	**-.08**	

Figure 2

Forest Plot for Meta-Analys



Blue circles represent the effect size of each study, numbers on the left-hand side correspond to the study numbers in Table 4. Study 8 represent the pooled combined effect size. Black horizontal lines represent confidence intervals, whilst green horizontal lines denote predictive intervals

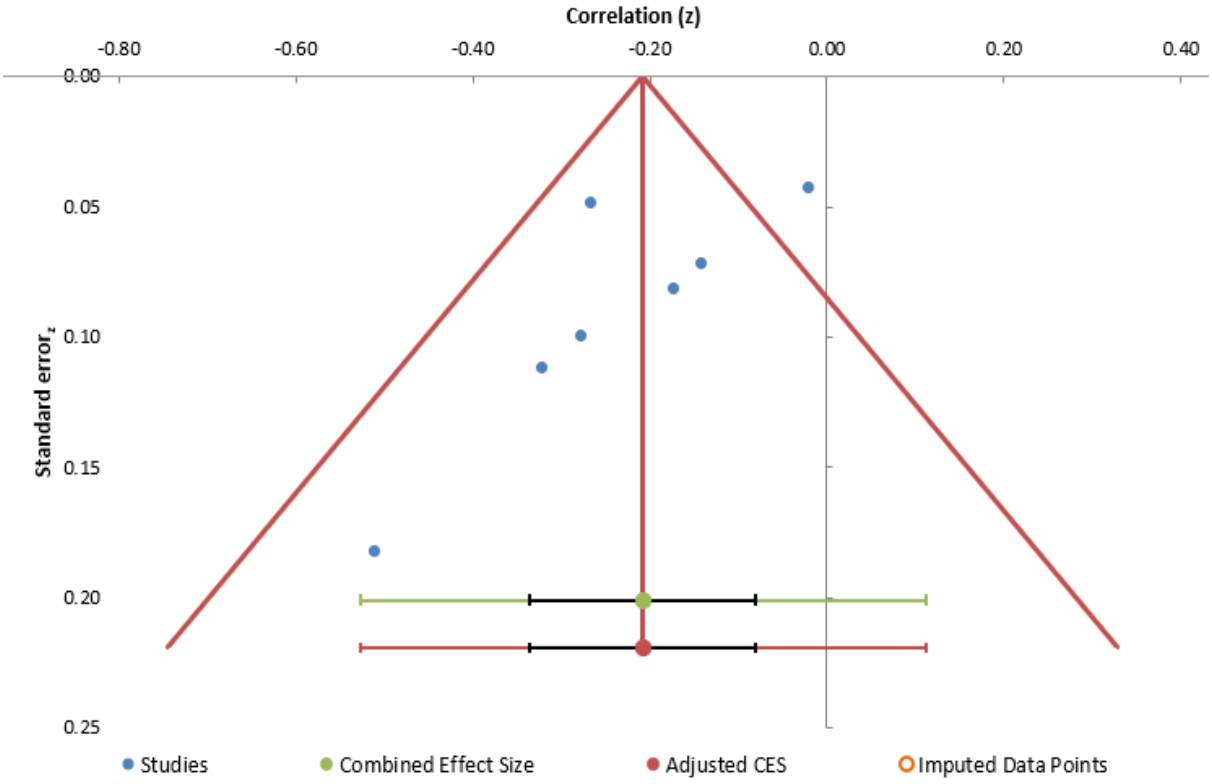
Publication bias

The fail-safe N analysis indicated that 34 studies with null results would be required to reduce the significance level of the effect and increase the p-value to >

0.05. This is below the threshold of 45 (calculated as 5k + 10, where k = number of studies) using guidance described by Rosenthal (1979), suggesting a possibility of publication bias. Egger's regression found a significant result: b_0 = -4.48 [-8.19, -0.76], *t*(13) = -2.94, p = .03, also indicating potential publication bias. However, the funnel plot (see Figure 3) was relatively symmetrical. However, this should be interpreted with caution, as the number of studies included in the meta-analysis was below the recommended threshold of ten, for reliably assessing publication bias using funnel plots (Lau et al., 2006). Nonetheless, the trim-and-fill test led to zero studies being trimmed, indicating low chance of publication bias.

Figure 3

Funnel Plot to Assess Publication Bias



Sensitivity Analysis

To explore if the observed effects were susceptible to influences by differences in how key variables were measures across studies, three sensitivity analyses were conducted. Specifically, exploring the effect of studies that utilised different measurement approaches for pre-treatment expectancy and depression outcomes, from the majority of included studies.

The potential impact of a studies that measured expectancy by using a single item measure instead of more comprehensive multi-item measures, as used in other studies was assessed by removing two studies one by one (independently). The effects were reviewed to determine whether their inclusion was impacting the initial observed effects. Firstly, the study by Høifødt et al. (2015) was identified and removed. Upon removal of this study the results showed minimal variation (r = - .19, 95% CI [-.33, -.04]; z = -3.34, p = .001), supporting the decision to retain this study in the analysis. Secondly, the study by Schindler et al. (2013) was removed and again results were largely unchanged (r = - .22, [-.37, -.07]; z =- 3.67, p < .001), again supporting to retain the study in the analysis.

The third sensitivity analysis was conducted to account for difference in both the depression outcome measurement and expectancy measurement. All studies relied on self-report depression measure, however the study by Vittengl et al. (2019), used a clinician rated Hamilton Rating Scale for Depression (HRSD) alongside a single-item measure of outcome expectancy. Clinician-rated measures can be seen as more objective and less susceptible to social desirability and can be influenced by clinician interpretation. Upon removal of this study the results showed minimal change (r = -.22, 95% CI [-.36, -.06]; z = -3.51, p = p < .001), supporting to retain the study in the analysis.

Narrative Summary

Three out of the four studies support the findings of the meta-analysis; a fourth study did not show a significant relationship between pre-treatment expectancies and treatment outcome. These studies are summarised below.

Studies Showing a Negative Association: Higher Outcome Expectancy Linked to Reduction in Depression Symptoms Post-Treatment

Three studies outlined that higher pre-treatment outcome expectancy was linked to reduction in depressive symptoms post-treatment. These implied that participants who had higher expectations for the effectiveness of treatment given, experienced a better outcome, with a greater decrease depressive symptom post-treatment in contrast to those with lower pre-treatment outcome expectancy.

Snippe et al. (2015) studied patients with diabetes and comorbid depression, receiving individual CBT. Outcome expectancy was assessed using a two-item questionnaire based on the work of Borkovec and Nau (1972) and Devilly and Borkovec (2000). The findings outlined that higher pre-treatment outcome expectancy was associated with reduced depressive symptoms post-treatment ($b = -2.2$, $p < .01$).

Freidl et al. (2020) studied the relationship between pre-treatment outcome expectancy and post-treatment outcome, in two treatment formats: blended CBT (individual standard CBT was combined with internet-based CBT elements) and traditional standard CBT (TAU). Outcome expectancy was measured using the Credibility/Expectancy Questionnaire. In the TAU group, the posterior probability was 97% (posterior mean: -0.27, SD: 0.11) for outcome expectancy predicting post-

treatment scores. In the blended treatment condition, the posterior probability was 72.4%, (posterior mean was -0.22, SD: 0.17).

Webb et al. (2012) measured pre-treatment expectancy using the CEQ, in a study involving group CBT therapy in combination with individual one-to-one sessions. Pre-treatment expectancy scores were submitted into an elastic net regression model and emerged as a predictor of post-treatment outcome (b = - 0.44), indicating higher pre-treatment outcome expectancy was associated with a reduction in depressive symptoms post treatment.

Study Showing No Clear Association between Pre-Treatment Outcome Expectancy and Reduction in Depression Symptoms Post-Treatment

Visla et al. (2018) studied outcome expectancy using the OES, in a study involving group CBT. The findings indicated that higher pre-treatment outcome expectancy was associated with slightly higher depression scores post-treatment, although this association was not statistically significant (b = 2.61, $p > .05$).

Discussion

This systematic review is, to the author's best knowledge, the first to examine the association between pre-treatment outcome expectancy and post-treatment outcomes in CBT for depression. A total of eleven studies were included within this review. Seven of these were included in the meta-analysis and four were narratively summarised as insufficient data was available for these studies to be included in the meta-analysis. In relation to the meta-analysis, a small to moderate negative association between pre-treatment outcome expectancy and post-treatment outcomes was found (r = - 0.21, 95% CI [-0.32, - 0.08]). This indicated that individuals who had higher expectations for treatment report lower depressive

symptoms following treatment. All studies within the meta-analysis indicated a negative correlation.

The narrative summary of the four studies provided further insightful context, reinforcing the negative association between pre-treatment outcome expectancy and post-treatment depression reduction. The majority of studies (N=3) within the narrative summary showed a negative relationship between pre-treatment outcome expectancy and post-treatment depression outcomes. Only one study by Visla et al. (2020) found a small positive relationship, which could be viewed as an outlier in an otherwise consistent trend. This study by Visla et al. (2018) was the only study out of the eleven that used the OES scale, whereas most studies within the review used the CEQ. Both the OES and CEQ are valid measure of pre-treatment outcome expectancy. However, the OES measures more general aspects of pre-treatment outcome expectancy, such as confidence in therapy, the expected success of the therapy and anticipated recovery from depression. Conversely, the CEQ specifically asks questions in relation to 'symptom improvement' focusing more precisely on how much one believes that that therapy will specifically reduce their symptoms. This measure is likely to provide a closer link to symptom improvement. Therefore, the use of the OES scale could potentially explain the differing results observed in the study by Visla et al. (2018).

It is important to note that some of the included studies controlled for baseline severity in their analyses, strengthening the validity of the observed associations. Overall, the findings of this review underscore the significance of pre-treatment beliefs shaping treatment outcome. The results are consistent with previous meta-analyses (Constantino et al., 2011; 2018), who also further that higher pre-treatment outcome expectancy to be associated with better treatment outcomes across various

therapeutic contexts and disorders. The findings of this review underpin the importance of several theories. Expectancy theory (Vroom, 1964) theorises one is more motivated and engaged when they expect their efforts to lead to positive outcomes. Findings also support the cognitive behavioural theory, which outlines the interplay between thoughts, feelings and behaviours, and how they influence each other (Dobson et al., 2019). Thus, patients with positive expectations may enter treatment with a more proactive mindset, enhancing their ability to engage with treatment and ultimately facilitating greater progress and benefits form therapy.

The findings of this review further show the relevance of the self-fulfilling prophecy theory (Merton, 1948) and the placebo effect (Benedetti, 2013). In relation to self-fulfilling prophecies, if one believes treatment will be effective in reducing their symptoms, they then may engage more actively. They are also more likely implement strategies learnt, and adopt an optimistic mindset, making the predicted outcome more likely to occur. Conversely negative expectations may lead to lack of effort and disengagement reinforcing poor outcomes. Lastly, placebo effect indicates that one may experience symptom improvement, simply because the recipient expects it to (Stewart Williams & Podd, 2004) even when the treatment is inert. While CBT is an active evidence-based treatment, pre-treatment outcome expectancy may function similarly to a placebo response, whereby positive expectations lead to greater perceived benefits.

The above theories suggest ways that pre-treatment outcome expectancies can influence post-treatment outcome. In clinical practice, strategies grounded in these theories, could be used to enhance pre-treatment outcome expectations (Constantino et al., 2012).

Methodological Consideration

A major strength of the review is its thorough search strategy. This included the use of broad search terms, searches across several databases, the inclusion of grey literature and backward and forward citation searching. This approach maximised the likelihood of identifying all relevant studies and reduced publication bias, strengthening the robustness of findings and ensuring a representative synthesis of the evidence. An additional strength is the use of the second rater during the full text screening process and the risk of bias assessment. This provides the review with inter-rater reliability, whilst reducing potential reviewer bias (McHugh, 2012) and increases the credibility and reliability of the results.

However, it should be noted that the search was limited to English language only, which may have led to the exclusion of appropriate studies published in other languages. This limitation could introduce language bias and limit the generalisability of the findings (Uttley et al., 2023). Whilst this can be seen as a limitation of the review, research has found that excluding non-English studies in systematic reviews has minimal effect on overall conclusions (Dobrescu et al., 2021; Nussbaumer-Streit, 2020). Nonetheless, future reviews would benefit from including non-English studies to increase the generalisability across various cultural and linguistic populations (Stern & Kleijnen, 2020). This would provide a more representative understanding of the topic.

A further limitation of this review is its solely focused on adult population only, reducing the generalisability of the findings to other age groups such as children and adolescents. Also, qualitative research was not included within the review. Inclusion of this could have provided valuable insights on the association of pre-treatment

outcome expectancy and post-treatment outcome, to complement the quantitative findings.

Quality appraisal was conducted using the CASP checklist. However, not all questions within the CASP checklists were relevant to the studies included, thus checklists were adapted to ensure relevance to the topic of interest. Certain questions were adapted or omitted to better align with the systematic review question. Such adaptations allowed for a more tailored appraisal of the studies but could have introduced some subjectivity into the process. CASP checklists do not include predefined scoring thresholds and to address this, custom thresholds were established to appraise the studies. This can introduce further subjectivity as thresholds were based on personal judgement rather than a standardised criterion. One may argue this could have reduced the reliability of the risk of bias assessment, as other reviewers may have applied different scoring thresholds.

Furthermore, this review examined the association between pre-treatment outcome expectancy and post-treatment outcome, thus causal inferences cannot be made (Papageorgiou, 2022). Therefore, findings should be interpreted with caution. Furthermore, by synthesizing correlations, it was not possible to control for baseline severity as a potential confounder across all studies. However, some individual studies controlled for baseline clinical and demographic factors and still found a significant association between outcome expectancy and treatment outcomes

Despite some guidelines generally recommending at least 10 studies for a robust publication bias assessment (Sterne et al., 2011), an assessment was still conducted to provide transparency. A strength is that publication bias was assessed using several methods. The trim-and-fill test showed no evidence of publication bias,

whilst the funnel plot was relatively symmetrical. However, Egger's regression test and the fail-safe N implied a potential risk of publication bias. Therefore, these findings should be interpreted with caution. Conducting a publication assessment on a meta-analysis with a small number of studies and high heterogeneity can lead to misleading results, specifically increasing the risk of false positives (Deeks et al., 2005). Nonetheless, a thorough search for grey literature was conducted and no additional unpublished studies were identified. Therefore, it could be implied overall when all is taken into consideration, the potential impact of publication bias is limited.

A moderate amount of heterogeneity (73.17%) was found within the meta-analysis, but there were insufficient studies to conduct moderator analyses and explore potential sources of heterogeneity (Card et al., 2012; Higgins et al., 2002; Pincus et al., 2011). Therefore, given the moderate amount of heterogeneity, findings should be interpreted with caution.

A limitation of the studies included within the review is that all utilised self-reported outcome expectancy measures and all but one utilised self-reported depression outcome measures. Using self-report measures bias can result in social desirability and response bias. This has the potential to inflate the correlation the correlation between pre-treatment expectations and post-treatment outcomes (Podsakoff, 2012). Future research could incorporate data from various sources, to mitigate self-report bias.

Some studies included in the review use one-item outcome expectancy measure, and often did not report the psychometric properties of these. Furthermore, it could be argued that a one-item expectancy measure may not fully capture the complexity of this construct. Therefore, sensitivity analyses were conducted to

further ensure robustises of the findings, and these indicated removing these studies from the analysis resulted in minimal change in the pooled effect size.

Studies included within the review utilised a variety of outcome expectancy and depression outcome measures. This can be viewed as both a strength and a weakness. On the one hand, pooling studies with various measure can introduces challenges and variation, as these measures assess a similar construct but by using different questions, units and direction. However, to account for some of this variability, a random-effects meta-analysis was employed. Conversely, including a variety of measures, allows one to capture a diverse range of studies, providing a comprehensive view of the relationship between outcome expectancy and post treatment outcomes. Given the limited number of studies within the review, solely focusing on one scale would have further limited the number studies, resulting in unclear estimates and wide confidence intervals (Murad et al., 2019). Therefore, it could be seen as adequate to include studies using different measures. If more studies become available, conducting subgroup analyses to explore potential variations across different types of scale, could provide a more nuanced understanding.

Clinical Implications

The findings of this review indicate that pre-treatment outcome expectancies play a role in determining the post-treatment depression outcomes in CBT. Therefore, it is important for clinicians to consider how they can address and manage pre-treatment outcome expectancy within their clinical practice.

Clinicians should assess outcome expectancy pre-treatment either through verbal discussion or by using a brief outcome expectancy measure. This would allow

clinicians to verify and validate these beliefs. As found within this review, patients with more optimist pre-treatment outcome expectations tend to experience better response to therapy, indicating the significance of understanding these expectations upfront. By recognising outcome expectancies early, therapist should identify and correct any unrealistic or overly negative expectations, alongside reinforcing realistic expectations.

Additionally, to enhance pre-treatment outcome expectancies, clinician could consider utilising motivational interviewing techniques (Miller & Rollnick, 2013) to address any ambivalence and reinforce positive outcome expectation about therapy. Such techniques can encourage more motivation, engagement and confidence (Rubak et al., 2005) with the aim of ultimately improving treatment outcomes. Constantino (2012) outlines how techniques such as motivation enhancement therapy and motivational interviewing can help to address ambivalence or resistance, address outcome expectancies and support self-efficacy,

Clinician could also address expectations with hope-inspiring statements, that neither abruptly challenge one's belief system or self-perception, nor set unrealistic expectations of change. Statements can be kept general such as "Your concerns are commonly seen in therapy and are exactly the kind therapy can assist with" (Constantino, et al., 2006).

To enhance pre-treatment outcome expectancies, clinician should provide clear and persuasive explanation of CBT before treatment commences. Clinicians can outline that CBT is a well-evidenced therapy for depression, which can address a wide range of emotional, cognitive, and behavioral factors. Furthermore, clinicians could also share examples of previous successful cases.

Conclusion

This review identified an association between pre-treatment outcome expectancy and post-treatment outcome in CBT for depression. Findings indicate individuals who enter therapy with higher expectations of the treatment are more likely to experience greater reductions in depressive symptoms. Acknowledging the limitations above, findings should be interpreted with caution.

Future Directions

Several avenues for future research remain. Firstly, future research should continue assessing the impact of pre-treatment outcome expectancy on outcome in CBT for depression to add to the small evidence base to date. From this, further replicated systematic reviews can be conducted to build on this current review.

Future research should conceptualise outcome expectations as something that can change over time, rather than a static construct. Expectancies could be measured at several time points, to see how this influences post-treatment outcome. Furthermore, more research is needed to explore the long-term impact of pre-treatment outcome expectancies on outcome, such as at 12 months after therapy. This would help to determine whether pre-treatment outcome expectancies influence not only post-treatment outcomes, but also the sustainability of progress and relapse rates over time. Finally, more research is needed to explore the impact and effectiveness of specific strategies aimed at enhancing outcome expectancies. This could provide important insights into how these strategies could potentially be incorporated.

References

Andresen, E. M., Malmgren, J. A., Carter, W. B., & Patrick, D. L. (1994). Screening for depression in well older adults: Evaluation of a short form of the CES-D enter for epidemiologic studies depression scale). *American Journal of Preventive Medicine*, *10*(2), 77–84.

Beard, C., Stein, A. T., Hearon, B. A., Lee, J., Hsu, K. J., & Björgvinsson, T. (2016). Predictors of depression treatment response in an intensive CBT partial hospital. *Journal of Clinical Psychology*, *72*(4), 297–310. https://doi.org/10.1002/jclp.22269

Beck, A. T., Steer, R. A., & Brown, G. (1996). Beck depression inventory–II. PsycTESTS Dataset. https://doi.org/10.1037/t00742-000

Benedetti F. (2013). Placebo and the new physiology of the doctor-patient relationship. *Physiological Reviews*, *93*(3), 1207–1246. https://doi.org/10.1152/physrev.00043.2012

Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry, 3*(4), 257–260. https://doi.org/10.1016/0005-7916(72)90045-6

Branson, A., Myles, P., Mahdi, M., & Shafran, R. (2018). The relationship between competence and patient outcome with low-intensity cognitive behavioural interventions. *Behavioural and Cognitive Psychotherapy*, *46*(1), 101–114. https://doi.org/10.1017/S1352465817000522

Card, N. A. (2012). *Applied meta-analysis for social science research*. Guilford Press.

Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge Academic.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Conner, K. R., Pinquart, M., & Gamble, S. A. (2009). Meta-analysis of depression and substance use among individuals with alcohol use disorders. *Journal of Substance Abuse Treatment*, *37*(2), 127–137. https://doi.org/10.1016/j.jsat.2008.11.007

Constantino M. J. (2012). Believing is seeing: An evolving research program on patients' psychotherapy expectations. *Psychotherapy Research*, *22*(2), 127–138. https://doi.org/10.1080/10503307.2012.663512

Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. *Journal of Clinical Psychology*, *67*(2), 184–192. https://doi.org/10.1002/jclp.20754

Constantino, M. J., Boswell, J. F., & Coyne, A. E. (2021). Patient, therapist, and relational factors. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change: 50th anniversary edition* (7th ed., pp. 225–262). John Wiley & Sons.

Constantino, M. J., Vîslă, A., Coyne, A. E., & Boswell, J. F. (2018). A meta-analysis of the association between patients' early treatment outcome expectation and their posttreatment outcomes. *Psychotherapy (Chicago, Ill.)*, *55*(4), 473–485. https://doi.org/10.1037/pst0000169

Constantino, M. J., Vîslă, A., Coyne, A. E., & Boswell, J. F. (2018). A meta-analysis of the association between patients' early treatment outcome expectation and their posttreatment outcomes. *Psychotherapy (Chicago, Ill.)*, *55*(4), 473–485. https://doi.org/10.1037/pst0000169

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

Deeks, J. J., Macaskill, P., & Irwig, L. (2005). The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of Clinical Epidemiology*, *58*(9), 882–893. https://doi.org/10.1016/j.jclinepi.2005.01.016

Delsignore, A., & Schnyder, U. (2007). Control expectancies as predictors of psychotherapy outcome: A systematic review. *British Journal of Clinical Psychology, 46*(4), 467–483. https://doi.org/10.1348/014466507X226953

Dettori, J. R., Norvell, D. C., & Chapman, J. R. (2022). Fixed-Effect vs random-effects models for meta-analysis: 3 points to consider. *Global Spine Journal*, *12*(7), 1624–1626. https://doi.org/10.1177/21925682221110527

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the Credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, *31*(2), 73–86. https://doi.org/10.1016/s0005-7916(00)00012-4

Dobrescu, A. I., Nussbaumer-Streit, B., Klerings, I., Wagner, G., Persad, E., Sommer, I., Herkner, H., & Gartlehner, G. (2021). Restricting evidence syntheses of interventions to English-language publications is a viable methodological

shortcut for most medical topics: A systematic review. *Journal of clinical Epidemiology*, *137*, 209–217. https://doi.org/10.1016/j.jclinepi.2021.04.012

Dozois, D. J. A., Dobson, K. S., & Rnic, K. (2019). Historical and philosophical bases of the cognitive-behavioral therapies. In K. S. Dobson & D. J. A. Dozois (Eds.), *Handbook of cognitive-behavioral therapies* (4th ed., pp. 3–31). The Guilford Press.

Duval, S., & Tweedie, R. (2000). A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*(449), 89-98. https://doi.org/10.1080/01621459.2000.10473905

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

Frank, J. D. (1961). *Persuasion and healing: A comparative study of psychotherapy.* Johns Hopkins University Press.

Friedl, N., Krieger, T., Chevreul, K., Hazo, J. B., Holtzmann, J., Hoogendoorn, M., Kleiboer, A., Mathiasen, K., Urech, A., Riper, H., & Berger, T. (2020). Using the personalized advantage index for individual treatment allocation to blended treatment or treatment as usual for depression in secondary care. *Journal of Clinical Medicine*, *9*(2), 490. https://doi.org/10.3390/jcm9020490

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: Current status and future directions. *Annual Review of Clinical Psychology*, *6*, 285–312. https://doi.org/10.1146/annurev.clinpsy.121208.131305

Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? *Clinical Psychology Review*, *26*(6), 657–678. https://doi.org/10.1016/j.cpr.2005.03.002

Hamilton, K. E., & Dobson, K. S. (2002). Cognitive therapy of depression: Pretreatment patient predictors of outcome. *Clinical Psychology Review*, *22*(6), 875–893. https://doi.org/10.1016/s0272-7358(02)00106-x

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*(1), 56–62. https://doi.org/10.1136/jnnp.23.1.56

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

Hardy, G. E., Barkham, M., Shapiro, D. A., Reynolds, S., Rees, A., & Stiles, W. B. (1995). Credibility and outcome of cognitive-behavioural and psychodynamic-interpersonal psychotherapy. *The British Journal of Clinical Psychology*, *34*(4), 555–569. https://doi.org/10.1111/j.2044-8260.1995.tb01489.x

Hautzinger, M., Bailer, M., Worall, H., & Keller, F. (1995). Beck-depressions-inventar. *Bern: Huber*.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*, *36*(5), 427–440. https://doi.org/10.1007/s10608-012-9476-1

Høifødt, R. S., Mittner, M., Lillevoll, K., Katla, S. K., Kolstrup, N., Eisemann, M., Friborg, O., & Waterloo, K. (2015). Predictors of response to web-based cognitive-behavioral therapy with high-intensity face-to-face therapist guidance for depression: A bayesian analysis. *Journal of Medical Internet Research*, *17*(9), e197. https://doi.org/10.2196/jmir.4351

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment, 8*(4), 275-292. https://doi.org/10.1111/1468-2389.00156

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139 193). Wiley.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lau, J., Ioannidis, J. P., Terrin, N., Schmid, C. H., & Olkin, I. (2006). The case of the misleading funnel plot. *BMJ, 333*(7568), 597–600. https://doi.org/10.1136/bmj.333.7568.597

Lenhard, W., & Lenhard, A. (2022). *Computation of effect sizes*. Psychometrica. https://doi.org/10.13140/RG.2.2.17823.92329

Lutz, W. (2002). Patient-focused psychotherapy research and individual treatment progress as scientific groundwork for an empirically based clinical practice. *Psychotherapy Research, 12*(3), 2, 51–272. https://doi.org/10.1093/ptr/12.3.251

McHugh M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

Merton, R. K. (1948). The self-fulfilling prophecy. *Antioch Review*, 8, 193–210.

Miller, W. R., & Rollnick, S. (2013). *Motivational interviewing: Helping people change*. Guilford Press.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264-269.

Murad, M. H., Wang, Z., Chu, H., & Lin, L. (2019). When continuous outcomes are measured using different scales: Guide for meta-analysis and interpretation. *BMJ (Clinical research ed.)*, *364*, k4817. https://doi.org/10.1136/bmj.k4817

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

Newman, M. G., & Fisher, A. J. (2010). Expectancy/credibility change as a mediator of cognitive behavioral therapy for generalized anxiety disorder: Mechanism of

action or proxy for symptom change? *International Journal of Cognitive Therapy*, *3*, 245–261. https://doi.org/10.1521/ijct.2010.3.3.245

Nussbaumer-Streit, B., Klerings, I., Dobrescu, A. I., Persad, E., Stevens, A., Garritty, C., Kamel, C., Affengruber, L., King, V. J., & Gartlehner, G. (2020). Excluding non-English publications from evidence-syntheses did not change conclusions: A meta-epidemiological study. *Journal of Clinical Epidemiology*, *118*, 42–54. https://doi.org/10.1016/j.jclinepi.2019.10.011

Ogrodniczuk, J. S., & Sochting, I. (2010). *Outcome expectancy scale*. The University of British Columbia.

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

Papageorgiou S. N. (2022). On correlation coefficients and their interpretation. *Journal of Orthodontics*, *49*(3), 359–361. https://doi.org/10.1177/14653125221076142

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61*(10), 991-996. https://doi.org/10.1016/j.jclinepi.2007.11.010

Pincus, T., Miles, C., Froud, R., Underwood, M., Carnes, D., & Taylor, S. J. (2011). Methodological criteria for the assessment of moderators in systematic reviews of randomised controlled trials: A consensus study. *BMC Medical Research Methodology*, *11*, 14. https://doi.org/10.1186/1471-2288-11-14

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63,* 539–569. https://doi.org/10.1146/annurev-psych-120710-100452

Ramondo, N., Pestell, C. F., Byrne, S. M., & Gignac, G. E. (2024). Cognitive behavioral therapy and hypnosis in the treatment of major depressive disorder: A randomized control trial. *The International Journal of Clinical and Experimental Hypnosis*, *72*(3), 229–253. https://doi.org/10.1080/00207144.2024.2354722

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641. https://psycnet.apa.org/doi/10.1037/0033-2909.86.3.638

Rubak, S., Sandbaek, A., Lauritzen, T., & Christensen, B. (2005). Motivational interviewing: A systematic review and meta-analysis. *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, *55*(513), 305–312.

Schindler, A., Hiller, W., & Witthöft, M. (2013). What predicts outcome, response, and drop-out in CBT of depressive adults? A naturalistic study. *Behavioural and Cognitive Psychotherapy*, *41*(3), 365–370. https://doi.org/10.1017/S1352465812001063

Schulte D. (2008). Patients' outcome expectancies and their impression of suitability as predictors of treatment outcome. *Psychotherapy Research: Journal of the Society for Psychotherapy Research*, *18*(4), 481–494. https://doi.org/10.1080/10503300801932505

Snippe, E., Schroevers, M. J., Annika Tovote, K., Sanderman, R., Emmelkamp, P. M., & Fleer, J. (2015). Patients' outcome expectations matter in psychological interventions for patients with diabetes and comorbid depressive symptoms. *Cognitive Therapy and Research*, *39*(3), 307–317. https://doi.org/10.1007/s10608-014-9667-z

Stern, C., & Kleijnen, J. (2020). Language bias in systematic reviews: You only get out what you put in. *JBI Evidence Synthesis*, *18*(9), 1818–1819. https://doi.org/10.11124/JBIES-20-00361

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ (Online),* 343(7818), 1–8. https://doi.org/10.1136/bmj.d4002

Stevinson, C., & Lawlor, D. A. (2004). Searching multiple databases for systematic reviews: Added value or diminishing returns?. *Complementary Therapies in Medicine*, *12*(4), 228-232. https://doi.org/10.1016/j.ctim.2004.09.003

Stewart-Williams, S., & Podd, J. (2004). The placebo effect: Dissolving the expectancy versus conditioning debate. *Psychological Bulletin*, *130*(2), 324–340. https://doi.org/10.1037/0033-2909.130.2.324

Suurmond, R., van Rhee, H., & Hak, T. (2017). Introduction, comparison, and validation of meta-essentials: A free and simple tool for meta-analysis. *Research Synthesis Methods*, *8*(4), 537–553. https://doi.org/10.1002/jrsm.1260

The EndNote Team (2013). *EndNote* (EndNote X9) [64 bit]. Clarivate.

Turecki, G., & Brent, D. A. (2016). Suicide and suicidal behaviour. *Lancet (London, England)*, *387*(10024), 1227–1239. https://doi.org/10.1016/S0140-6736(15)00234-2

Tylee, A., & Gandhi, P. (2005). The importance of somatic symptoms in depression in primary care. *Primary Care Companion to the Journal of Clinical Psychiatry*, *7*(4), 167–176. https://doi.org/10.4088/pcc.v07n0405

Uttley, L., Quintana, D. S., Montgomery, P., Carroll, C., Page, M. J., Falzon, L., Sutton, A., & Moher, D. (2023). The problems with systematic reviews: A living systematic review. *Journal of Clinical Epidemiology*, *156*, 30–41. https://doi.org/10.1016/j.jclinepi.2023.01.011

Vîslă, A., Constantino, M. J., Newkirk, K., Ogrodniczuk, J. S., & Söchting, I. (2018). The relation between outcome expectation, therapeutic alliance, and outcome among depressed patients in group cognitive-behavioral therapy. *Psychotherapy Research, 28*(3), 446–456. https://doi.org/10.1080/10503307.2016.1218089

Vittengl, J. R., Clark, L. A., Thase, M. E., & Jarrett, R. B. (2019). Could treatment matching patients' beliefs about depression improve outcomes? *Behavior Therapy*, *50*(4), 765–777. https://doi.org/10.1016/j.beth.2018.11.007

Vroom, V. H. (1964). *Work and motivation*. Wiley.

Webb, C. A., Beard, C., Auerbach, R. P., Menninger, E., & Björgvinsson, T. (2014). The therapeutic alliance in a naturalistic psychiatric setting: Temporal relations with depressive symptom change. *Behaviour Research and Therapy*, *61*, 70–77. https://doi.org/10.1016/j.brat.2014.07.015

Webb, C. A., Cohen, Z. D., Beard, C., Forgeard, M., Peckham, A. D., & Björgvinsson, T. (2020). Personalized prognostic prediction of treatment outcome for depressed patients in a naturalistic psychiatric hospital setting: A comparison of machine learning approaches. *Journal of Consulting and Clinical Psychology*, *88*(1), 25–38. https://doi.org/10.1037/ccp0000451

Webb, C. A., Kertz, S. J., Bigda-Peyton, J. S., & Björgvinsson, T. (2013). The role of pretreatment outcome expectancies and cognitive-behavioral skills in symptom improvement in an acute psychiatric setting. *Journal of Affective Disorders*, *149*(1-3), 375–382. https://doi.org/10.1016/j.jad.2013.02.016

Werson, A. D., Meiser-Stedman, R., & Laidlaw, K. (2022). A meta-analysis of CBT efficacy for depression comparing adults and older adults. *Journal of Affective Disorders*, *319*, 189–201. https://doi.org/10.1016/j.jad.2022.09.020

Win, S., Parakh, K., Eze-Nliam, C. M., Gottdiener, J. S., Kop, W. J., & Ziegelstein, R. C. (2011). Depressive symptoms, physical inactivity and risk of cardiovascular mortality in older adults: The cardiovascular health study. *Heart (British Cardiac Society)*, *97*(6), 500–505. https://doi.org/10.1136/hrt.2010.209767

World Health Organization. (2023). Depressive Disorder (depression). World Health Organization. https://www.who.int/news-room/fact-sheets/detail/depression [accessed 13/03/2025]

Appendices

Appendix A: PRISMA 2020 Guidelines

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Page 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	See below
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Page 4-8
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Page 7
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Page 12 & 16
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.	Pages 8-10
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Pages 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.	Pages 11-14
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	Page 14

Section and Topic	Item #	Checklist item	Location where item is reported
		applicable, details of automation tools used in the process.	
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.	Page 17
	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.	Page 14
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.	Pages 15-16
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.	Page 17
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)).	N/A
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Page 17
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Page 17 & 19
	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.	Pages 16-19
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	N/A
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Page 30
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	N/A
Certainty	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	N/A

Section and Topic	Item #	Checklist item	Location where item is reported
assessment			
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.	Page 19 & Page 22
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Page 20
Study characteristics	17	Cite each included study and present its characteristics.	Page 21 & 24
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Page 25 & Page 70-71
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.	Pages 27-28 & Pages 31-32
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	N/A
	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.	Pages 26-32
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Page 18
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Page 29-30
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	N/A
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	N/A

Section and Topic	Item #	Checklist item	Location where item is reported
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Pages 32-33
	23b	Discuss any limitations of the evidence included in the review.	Pages 35-38
	23c	Discuss any limitations of the review processes used.	Pages 35-38
	23d	Discuss implications of the results for practice, policy, and future research.	Pages 38-39 & Page 40
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.	Page 8
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Page 8
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N/A
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	N/A
Competing interests	26	Declare any competing interests of review authors.	N/A
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.	N/A

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.	YES-INCLUSION CRITERIA ONLY
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			
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies.	YES

Topic	No.	Item	Reported?
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.	N/A
Registration	12	Provide the register name and registration number.	YES

Appendix B: CASP Cohort Checklist

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle Way Oxford OX2 7LG

CASP Checklist: 12 questions to help you make sense of a **Cohort Study**

How to use this appraisal tool: Three broad issues need to be considered when appraising a cohort study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 12 questions on the following pages are designed to help you think about these issues systematically. The first two questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Cohort Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

CASP
Critical Appraisal Skills Programme

Paper for appraisal and reference:.....

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?

CASP
Critical Appraisal Skills Programme

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 ☐

CASP
Critical Appraisal
Skills Programme

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:

CASP
Critical Appraisal Skills Programme

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 C: CASP RCT Checklist

CASP
Critical Appraisal
Skills Programme

CASP Randomised Controlled Trial Standard Checklist:
11 questions to help you make sense of a randomised controlled trial (RCT)

Main issues for consideration: Several aspects need to be considered when appraising a randomised controlled trial:

- Is the basic study design valid for a randomised controlled trial? (Section A)
- Was the study methodologically sound? (Section B)
- What are the results? (Section C)
- Will the results help locally? (Section D)

The 11 questions in the checklist are designed to help you think about these aspects systematically.

How to use this appraisal tool: The first three questions (Section A) are screening questions about the validity of the basic study design and can be answered quickly. If, in light of your responses to Section A, you think the study design is valid, continue to Section B to assess whether the study was methodologically sound and if it is worth continuing with the appraisal by answering the remaining questions in Sections C and D.

Record 'Yes', 'No' or 'Can't tell' in response to the questions. Prompts below all but one of the questions highlight the issues it is important to consider. Record the reasons for your answers in the space provided. As CASP checklists were designed to be used as educational/teaching tools in a workshop setting, we do not recommend using a scoring system.

About CASP Checklists: The CASP RCT checklist was originally based on JAMA Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL and Cook DJ), and piloted with healthcare practitioners. This version has been updated taking into account the CONSORT 2010 guideline (http://www.consort-statement.org/consort-2010, accessed 16 September 2020).

Citation: CASP recommends using the Harvard style, i.e. *Critical Appraisal Skills Programme (2020). CASP (insert name of checklist i.e. Randomised Controlled Trial) Checklist. [online] Available at: insert URL. Accessed: insert date accessed.*

Study and citation: .....

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 ☐

SP
praisal
s Programme

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 ☐

Skills Programme

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 D: Adapted CASP checklist and Final Quality Ratings

CASP Cohort Checklist: Quality Appraisal

Study	Clear Focussed Issue?	Cohort recruited Acceptably ?	Accuracy of OE Measure?	Accuracy of Outcome Measure (Depression) ?	Identification of Confounding Variables?	Accounted for Confounding Variables?	Completeness of Results?	Precision of Results Reported ?	Believability of Results?	Applicability to Local Population?	Results Fit Wider Evidence?	Implications for Practice?	Overall Quality Rating
Beard et al. (2016)	Y	Y	Y	Y	Y	CT	Y	Y	Y	CT	Y	Y	High: 10/12:
Friedl et al. (2020)	Y	Y	CT	Y	Y	CT	Y	Y	Y	Y	Y	CT	Moderate 9/12
Høifødt et al. (2015)	Y	Y	CT	Y	CT	CT	Y	Y	CT	Y	CT	N	Moderate 6/12
Schindler et al. (2013)	Y	Y	CT	Y	N	CT	CT	N	CT	Y	CT	Y	Low 5/12
Snippe et al. (2015)	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	High 11/12
Vittengl et al. (2019)	Y	Y	CT	Y	Y	CT	Y	N	Y	Y	CT	N	Moderate 7/12
Visla et al. (2020)	Y	CT	Y	Y	Y	Y	Y	N	Y	Y	N	CT	Moderate 8/12
Webb et al. (2013)	Y	Y	Y	Y	Y	CT	Y	Y	Y	CT	Y	Y	High 10/12
Webb et al. (2014)	Y	Y	Y	Y	Y	CT	Y	N	Y	CT	Y	Y	Moderate 9/12
Webb et al. (2020)	Y	Y	CT	Y	CT	CT	Y	CT	CT	CT	Y	Y	Moderate 6/12

CASP RCT Checklist: Quality Appraisal

Study	Clearly focussed research question	Accuray of OE measure	Accuracy of Depression Measure	Are all participants accounted for	Are study groups similar?	Similar care across groups (except for experimental intervention)?	Effects of OE on outcome reported comprehensively?	Precision of OE and outcome reported	Results applicable to local population/ in your context?	Overall Quality Rating
Ramondo et al. (2024)*	**Y**	**CT**	**Y**	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**Moderate 7/9**

Appendix E: Reasons for Study Exclusion at Full Text Screening

Paper No.	Paper Title	Author	Full Text Screening Exclusion Reason	Summarised Reason
1	Predicting outcome in internet-based cognitive behaviour therapy for major depression: A large cohort study of adult patients in routine psychiatric care	Alaoui et al. (2016)	Focuses on treatment credibility instead of outcome expectancy	Does not explore topic of interest
2	Early therapy interpersonal process differentiating clients high and low in outcome expectations	Ahmed et al. (2012)	Focuses on GAD and therapist-client dynamics, not directly on pre-treatment outcome expectancy and outcome	Does not explore topic of interest
3	A Mobile-Based Intervention to Increase Self-esteem in Students With Depressive Symptoms: Randomized Controlled Trial	Bruhns et al. (2021)	Includes several therapy modalities in one therapy/not just CBT/no meaningful therapist contact	Not intervention of interest/No meaninfgul therapist involvement
4	Personality, expectancies and group psychotherapy	Caine et al. (1976)	Not specific to depression and CBT; focusses on neurotic patients in general	Does not explore topic of interest
5	Predictors of Adherence in Three Low-Intensity Intervention Programs Applied by ICTs for Depression in Primary Care	Castro et al. (2021)	Does not focus on topic of interest as looks at adherance instead of treatment outcome	Does not explore topic of interest
6	The acceptability of computer-aided cognitive behavioural therapy: A pragmatic study	Cavanagh et al. (2009)	No direct therapist involvement/study population = anxiety and depression; separate outcome for anxiety not reported	No meaningful therapist involvement
7	Cognitive-Behavioral Depression Treatment for Mothers of Children With Attention-Deficit/Hyperactivity Disorder	Chronis et al. (2006)	Does not assess outcome expectancy or its association on treatment outcomes in CBT for depression	Does not explore topic of interest

8	Credibility and outcome of cognitive-behavioural and psychodynamic-interpersonal psychotherapy	Hardy et al. (1995)	Does not explore topic of interest - Looks at credibility, not outcome expectancy	Does not explore topic of interest
9	The prediction of psychotherapy success by outcome expectations in inpatient psychotherapy	Holtforth et al. (2011)	Does not focus on depression participants only – sample is various disorders and results for depressive sample not outlined. Says Affective disorders but not clear which diagnosis come under this	Not appropriate clinical sample
10	Mental health treatment outcome expectancies in Burundi	Irankunda et al. (2017)	Does look at outcome expectancies in general, but not in association with how this impacts treatment outcome	Does not explore topic of interest
11	Goal Achievement and Goal-Related Cognitions in Behavioral Activation Treatment for Depression	Knittle et al. (2019)	Focusses on goal-specific outcome expectancy, not general treatment outcome expectancy	Does not explore topic of interest
12	Development and initial evaluation of blended cognitive behavioural treatment for major depression in routine specialized mental health care	Kooistra et al. (2016)	Does not explore topic of interest (outcome expectancy and treatment outcome)	Does not explore topic of interest
13	Patient Choice in Depression Psychotherapy Outcomes of Patient-Preferred Therapy Versus Randomly Allocated Therapy	Kuzminskaite et al. (2021)	Study combines multiple therapy modalities without separate data for Cognitive Therapy (CT)/focusses on preference	Does not explore topic of interest
14	Treatment expectations for cognitive-behavioral therapy and light therapy for seasonal affective disorder: Change across treatment and relation to outcome	Meyerhoff et al.(2016)	Does not look at pre-treatment outcome expectancy and outcome; instead looks whether change (throughout sessions) in treatment expectations predicts outcome	Outcome expectancy not appropriately assessed
15	Increasing Acceptability and Outcome Expectancy for Internet-Based Cognitive Behavioral Therapy During the COVID-19 Pandemic	Molloy et al. (2022)	Does not explore topic of interest (outcome expectancy and treatment outcome)	Does not explore topic of interest
16	Expectancy and outcome in prescriptive vs. exploratory psychotherapy	Morrison et al. (1987)	Focusses on credibility not on outcome expectancy	Does not explore topic of interest
17	The acceptability and efficacy of a group cognitive behavioural therapy programme in a community mental health setting	Naik et al. (2013)	Does not explore topic of interest (outcome expectancy and treatment outcome)	Does not explore topic of interest
18	The association between positive outcome expectancies and avoidance in predicting the outcome of cognitive behavioural therapy for major depressive disorder	Renaud et al. (2013)	Combines client and clinician outcome expectancy scores to look at association between outcome expectancy and treatment outocme	Outcome expectancy not appropriately assessed
19	Patients' outcome expectancies and their impression of suitability as predictors of treatment outcome	Schulte et al. (2008)	Measures outcome expectancy at session 4; this is not pre-treatment	Outcome expectancy not appropriately assessed
20	Impact of baseline characteristics on the effectiveness of disorder-specific Cognitive Behavioral Analysis System of Psychotherapy (CBASP) and supportive psychotherapy in outpatient treatment for persistent depressive disorder	Serbanescu et al. (2020)	Does not focus on intervention of interest which is CBT, instead focusses CBASP	Not intervention of interest
21	Increasing hope by addressing clients' outcome expectations	Swift et al. (2013)	The study centres on a conceptual framework, not empirical evidence which would be relevant to subject of review.	Does not explore topic of interest
22	Does outcome expectancy predict outcomes in online depression prevention? Secondary analysis of randomised-controlled trials	Thielecke et al. (2024)	No meaningful therapist involvement	No meaningful therapist contact
23	The role of outcome expectancy in therapeutic change across psychotherapy versus pharmacotherapy for depression	Thiruchselvam et al. (2019)	Used DCES which assesses both broad as well as treatment-focused expectations for change in depression symptom. Therefore, measure not solely focussing on outcome expectancy	Outcome expectancy not appropriately assessed
24	Credibility and outcome expectancy in the unified protocol: Relationship to outcomes	Thompson-Hollands et al. (2014)	Principal primary diagnosis is anxiety not depression	Not appropriate clinical sample
25	Forecasting success: patients' expectations for improvement and their relations to baseline, process and outcome variables in group cognitive-behavioural therapy for depression	Tsai et al. (2014)	Does not report BDI outcome - reported BAI outcome which isn't appropriate	Appropriate outcome not reported

26	Predictors of treatment change and engagement in cognitive-behavioral group therapy for depression	Westra et al. (2022)	Hopelessness scaled used which was too broad; does not solely focus on outcome expectancy	Outcome expectancy not appropriately assessed
27	Feasibility, Acceptability, and Preliminary Efficacy of a Smartphone App–Led Cognitive Behavioral Therapy for Depression Under Therapist Supervision:	Wilhelm et al. (2024)	Does not explore topic of interest (outcome expectancy and treatment outcome)	Does not explore topic of interest
28	Factors contributing to symptom change in standardized and individualized Internet-based interventions for depression	Zagorscak et al. (2020)	No meaningful contact with therapist	No meaningful therapist involvement

Part Two: Empirical Study

Understanding the Role of Pre-Treatment Outcome Expectations, Treatment Preference and Wait Times on Non-attendance at the First Treatment Appointment and Clinical Outcomes in Person-Centered Experiential Therapy (PCET) and Cognitive Behavioural Therapy (CBT) for Depression: A Secondary Analysis of the PRaCTICED Trial Data Set

Abstract

Objectives: Depression is a global health concern, with a proportion of many individuals not benefiting from psychological therapy. Drop-out in mental health settings presents a further challenge. The study aimed to investigate whether pre-treatment factors such as pre-treatment outcome expectancy, treatment preference, and wait times, predicted non-attendance before therapy commenced and influenced clinical outcomes. A secondary aim was to explore whether treatment modality (cognitive behavioural therapy vs person-centred experiential therapy) moderated any significant relationships.

Design: A secondary data analysis utilising data from a single pragmatic randomised controlled trial.

Method: Data was derived from the PRaCTICE trial (N=510). Eight participants who switched treatment at baseline (did not receive allocated treatment) were excluded. After applying the inclusion criteria, the final sample for the non-attendance analysis was N=502, and for clinical outcomes N=369. Logistic and multiple regression analyses were conducted to study the associations between outcome expectancy, treatment preference and wait times with non-attendance and clinical outcomes. Non-attendance was defined as participants who ‘dropped-out’ after screening, before the first session. Clinical outcome was defined as the final session score on the standardised outcome measure PHQ-9. The association between wait times and non-attendance was not studied, due to wait time data not being available for individuals who dropped out.

Results: Higher pre-treatment outcome expectancy significantly predicted both lower PHQ-9 scores at post-treatment and a greater likelihood of non-attendance before

therapy commenced. Treatment preference and wait times did not predict non-attendance or clinical outcomes. Therapy modality did not moderate the significant relationships.

Conclusion: Findings suggest that outcome expectancy plays a complex role in therapy for depression, in relation to drop out and therapeutic outcomes. Higher outcome expectations were linked to a higher likelihood of non-attendance, but better clinical outcomes.

Key Practitioner Points:

- Higher pre-treatment outcome expectancy was associated with lower post-treatment PHQ-9 scores. However, high pre-treatment outcome expectancy also increased the likelihood of non-attendance before therapy began.
- Clinicians should assess outcome expectancies prior to treatment commencing.
- Outcome expectations should be carefully managed through strategies such as motivational interviewing and setting realistic expectations. Such strategies can optimise clinical outcomes whilst reducing the risk of non-attendance.

Key words: Depression, psychological therapies, outcome expectancy, wait times, treatment preference, non-attendance, clinical outcomes

Introduction

Depression is a multifaceted mental health disorder, and a major public health issue in the United Kingdom and globally (Norman & Ryrie, 2018). Symptoms differ across individuals, but typically include sadness, hopelessness, social withdrawal, lack of motivation and changes in sleep and appetite (Kanter, 2008; Kennedy, 2008). Over 280 million individuals worldwide suffer from depression, and it has been recognised as the leading cause of disability, accounting for 4.3% of the global disease burden (WHO, 2023).

Psychological therapies have been found to be broadly effective in the treatment of various mental health disorders, including depression (Bortolotti et al., 2008; Munder et al., 2019). However, 15% - 45% of clients show no clinically significant improvement following treatment (Hansen et al., 2002), and up to 10% deteriorate (Lambert & Ogles, 2004). Furthermore, mental health settings also report higher non-attendance rates, with 20% - 50% of clients failing to follow through after taking the initial steps towards treatment (Filippidou, 2014; Garfield, 1994). Identifying and supporting clients at risk of poor outcomes remains central to patient-focused research (Lutz, 2002).

Personalised psychological treatments aim to optimise treatment outcome by tailoring treatment to specific needs, preferences or other characteristics, through systematic adaptation of treatment or a differentiation between treatment strategies (Harnas et al., 2024). A meta-analysis of nine studies found enhanced outcomes for patients receiving personalised treatment versus standardised therapy (Nye et al., 2023). Studies in this review included only demographic and symptom measures to model personalised treatments for depression and anxiety. Other pre-treatment

factors that might potentially improve such models include what patients expect regarding their treatment outcomes, and their treatment preferences. Service characteristics such as the length of time they wait before treatment begins, may also play a role.

Outcome expectancy refers to a client's prognosis of how they will respond to or benefit from a planned treatment (Constantino et al., 2011). This differs from treatment credibility, which refers to a clients' understanding of how logical, convincing and believable a treatment is (Constantino, et al., 2011; Devilly & Borkovec, 2000; Greenberg et al., 2006).

A meta-analysis by Constantino et al. (2018) found a small positive association between pre- or early treatment outcome expectancy and clinical outcomes across various therapies and disorders. Most existing research has reported a significant relationship, with higher expectations associated with better treatment responses (Delgadilo et al., 2016; Price et al., 2012; Webb et al., 2013). Whilst a few studies have not found an association (Visla et al., 2018; Vogel, et al., 2006), studies specifically focusing on outcome expectancy and treatment outcome in depression are limited.

In addition, few studies have explored the link between outcome expectancy and treatment attendance and drop-out, specifically non-attendance at the first therapy session after initial screening. Existing studies have not found a clear relationship between pre-treatment expectations and initial therapy attendance. For example, Swift et al. (2012) found no significant relationship between pre-treatment outcome expectancy and attendance at an initial appointment after a brief screening. Norberg et al. (2011) found clients who did not attend the first therapy session had equal outcome expectations compared to those who attended. Recently, Bowker et

al. (2022) found that clients who believed therapy would be helpful were more likely to start treatment, but these initial outcome expectations were not associated with continued engagement.

These studies contain limitations: Norberg et al. (2011) did not control for confounders, such as symptom severity and demographics when considering the role of outcome expectancy on treatment attendance (Hanevic et al., 2023; Lester & Harris, 2007; Morton, 1995). Swift et al. (2012) had a small sample size and Bowker et al. (2022) used a single question to assess outcome expectancy. This could potentially be less reliable than multiple-item tools such as the Credibility/ Expectancy Questionnaire (CEQ; Devilly & Borkovec, 2000).

Patient preference refers to the therapist or therapy characteristics a client appreciates or desires (Glass et al., 2001). Findings on whether receiving a preferred treatment improves outcomes are mixed. A meta-analysis of 53 in-person therapy trials, with a sample of over 16,000 clients revealed that those assigned their choice of therapy had lower drop-out rates and better clinical outcomes (Swift et al., 2018). A later meta-analysis by Windle (2020), entailing 5294 participants, showed that receiving preferred psychosocial treatment reduced drop-out, but it had no effect on clinical outcomes. Davis et al. (2023) found patients who chose their treatment showed less adherence and smaller symptom reductions than those randomly assigned to acceptance or commitment therapy or cognitive behavioural therapy (CBT).

Despite these mixed findings, there are ethical and good practice reasons for involving patients in decision making and treatment choice (Slade, 2017). For example, the National Institute for Health and Care Excellence (NICE, 2012) outline that patients' treatment preferences should be considered when clinicians decide the

most appropriate type of treatment. The American Psychological Association (APA), alongside other healthcare organisations, recognises incorporating client preferences as vital to best practice standards (APA, 2006; Institute of Medicine, 2001).

Finally, treatment outcomes may be impacted by the length of time patients wait before beginning treatment. In 2022/23, nearly five million patients sought support from NHS mental health services, an increase of over a million from 2016/2017 (NHS England, 2024). Mental health services frequently experience prolonged waiting periods to offer psychological interventions such as counselling and CBT (Punton et al., 2022). Waiting lists can help to manage high demand (Culyer & Cullis, 1976). However, wait times in mental health services are known to be ‘too long’ for those with mental health difficulties, given their heightened vulnerability and need for immediate care (Smith et al., 2018). Long wait times may impede the benefit a patient gains from the treatment.

Several studies in the field of physical health have explored the impact of wait times on treatment outcomes (Braybrooke et al., 2007; Hirvonen et al., 2009; Tuominen et al., 2009). However, few studies have examined how waiting for therapy affects clinical outcomes in mental health services, specifically for depression. Van Dijk et al. (2023) found that prolonged waiting times were associated with poor clinical outcomes at six months for patients with depression. Reichert and Jacobs (2018) found a significant association between longer waiting times and worsened outcomes in psychosis, particularly when wait times exceed three months. However, some research has found minimal impact of wait times on clinical outcomes (Beck, 2015). Whilst other research has suggested natural remission before treatment,

thereby reducing the need for intervention after the waiting period (Whiteford et al., 2013).

The generalisability of these findings is limited as Reichert and Jacobs (2018) focused solely on psychosis. It further used clinician-reported outcome measures that may not match the client's perception (Kramer et al., 2003), whilst facing issues with missing data. Van Dijk et al.'s (2023) study comprised a small sample, and utilised an uncontrolled observational design, thereby reducing validity. Van Dijk et al. (2023) did study clients with depression, but they did not detail which treatments were offered, only noting that clients received pharmacotherapy and/or psychotherapy. Finally, previous research has mainly examined the link between wait time and clinical outcome, at follow-up, rather than post-treatment.

Clinical Implications

Overall, there is a clear need to further our understanding of predictors of non-attendance and therapeutic outcomes, specifically in a population group of clients experiencing depression, a debilitating condition. Mental health services face increasing pressure regarding resources. An enhanced understanding of the factors associated with non-attendance and clinical outcomes for patients could improve efficiency, increased use of services, and financial savings (Windle et al., 2020).

Study Aims and Hypotheses

Therefore, this study aimed to explore how treatment preference, outcome expectancy, and wait times impact non-attendance and clinical outcomes in clients with major depressive disorder, treated with CBT or person-centred experiential therapy (PCET). It further investigated the potential interaction effect between these factors and therapy type, as to the author's knowledge no literature has addressed this effect.

The overall aim is to conduct a secondary analysis of routine data assessing the impact of waiting times and patient preference and outcome expectancies on two outcomes: non-attendance at the first session of therapy (after screening) and post-treatment depression scores.

Accordingly, the study will address the two research questions:

1) Does pre-treatment outcome expectancy and treatment preference predict non-attendance at first session of therapy, and are the effects moderated by therapy type?
2) Does outcome expectancy, treatment preference, and wait times predict end of therapy scores and is this moderated by therapy type?

Based on previous literature it was hypothesised that:

1) Clients with higher pre-treatment outcome expectancies will be more likely to attend the first therapy session (less likely to drop-out before starting therapy).
2) Clients who receive their preferred therapy (CBT or PCET) will be more likely to attend the first therapy session (i.e., less likely to drop-out before starting therapy).
3) Clients with higher pre-treatment outcome expectancies will have significantly lower depression scores at the end of therapy.
4) There will be an association between clients receiving their preferred therapy (CBT or PCET) and depression scores at the end of therapy.
5) Clients who wait longer to access therapy will have significantly higher depression scores at the end of the therapy.

Method

The study has been completed in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE; Elm et al., 2007) guidelines (see appendix A).

Design

This current study utilised existing data from a pragmatic randomised non-inferiority trial comparing the clinical and cost effectiveness of PCET and CBT for the treatment of moderate or severe depression (PRaCTICED; Barkham et al., 2021). This data was used to conduct regression analyses to determine significant predictors of non-attendance at the first treatment appointment (before therapy began) and clinical outcome. This study was registered on AsPredicted (reference number: 185739).

Ethical Approval

Ethical approval for the PRaCTICED trial was obtained by the Health Research Authority (Research Ethics Committee 14/YH/0001; Barkham et al., 2021). To conduct these secondary data analyses, ethical approval was granted from The University of Sheffield Research Ethics Committee (12/01/2024; Reference: 058286, Appendix B). Participants provided informed consent, and this secondary research was conducted within the scope of the ethical approval.

PRaCTICED Trial Overview

The PRaCTICED trial was a pragmatic, randomised, non-inferiority trial, comparing the efficacy and cost effectiveness of PCET and CBT. The trial was embedded within a local primary care Improving Access to Psychological Therapy (IAPT) service, now known as NHS Talking Therapies for Anxiety and Depression service. The service utilised a stepped-care approach, offering a range of low

intensity and high intensity psychological therapies for the treatment of common mental health difficulties (Clark, 2011). The service served a population of approximately 560,000, and CBT and PCET were the two main high intensity psychological therapies offered.

PRaCTICED Trial Recruitment Process

The PRaCTICED trial followed a two-part recruitment process. In line with the service's usual practice, all clients were initially seen and assessed by a low-intensity Psychological Wellbeing Practitioner (PWP). Clients were introduced to the trial if 1) their main presenting concern was depression, 2) they had a PHQ-9 score of ≥ 12, and (3) they expressed no strong preference for either CBT or PCET (the first stage). This was checked by the PWPs, at this first stage and then verified at the screening interview. Those who were not willing to be randomised. continued with treatment as usual. If clients expressed an interest in the trial, they were provided with an information sheet and a consent-to-contact form. On returning this form, clients were invited to a screening interview (the second stage) with either the Clinical Support Officers or with a member of the research team. At screening, clients were asked again whether they had a strong preference for either treatment, to the extent that they would refuse a specific treatment if offered against their clear preference.

During the second stage screening clients completed the Clinical Interview Schedule-Revised (CIS-R), to determine whether they met the criteria for moderate or severe depression. Participants were randomised to either PCET or CBT via a remote-web based system managed by a body independent of the trial.

Trial Participants

In the PRaCTICED trial, 632 patients were assessed for eligibility and a final sample of 510 patients were randomly assigned to either PCET (254) or CBT (256).

Trial Patients

Participants had to be at least 18 years of age, and willing to work on their depression. Individuals with an organic condition, previous diagnoses of personality disorder, schizophrenia, or bipolar, were alcohol or substance dependent, or had high risk of suicide were further excluded. If one had been taking medication for depression, they had to have been on a stable regimen for the previous 6 weeks. If deemed eligible for the trial, clients signed a second stage consent form regarding randomisation and treatment allocation.

Trial Therapists

Trial CBT therapists and PCET counsellors had a comparable level of professional training, meeting the standards of their respective professional bodies and were eligible to be IAPT high-intensity practitioners. PCET was delivered by 18 PCET trained counsellors. Expert trainers assessed the PCET counsellor's adherence to the model by assessing recordings and using the 10-item Person-Centred Experiential Psychotherapy Scale (PCEPS). Only those who met the training requirements were eligible to deliver the therapy in the trial. In the PRaCTICED trial a standardised approach was followed to deliver PCET, by using the manual based on the book 'Counselling for depression: a person-centred and experiential approach to practice" (Sanders & Hill, 2014). On average the counsellors had been working within their current role for 9.6 years (SD = 5.46) and had on average 4.5 years of experience of delivering PCET. CBT was delivered by 32 CBT therapists. All CBT therapists received Beckian CBT refresher training from the research team prior to participating in the trial. Therapist's adherence to the

model was assessed using the 12-item Cognitive Therapy Scale-Revised (CTS-R). A treatment manual for CBT was developed and used by clinicians. On average the therapists had 8.1 (SD = 3.42) years of professional experience. Therapists received both individual and group supervision. Clients were offered a maximum of 20 sessions in either intervention. Once randomised they received therapy as standardly delivered within the service.

Trial Therapies

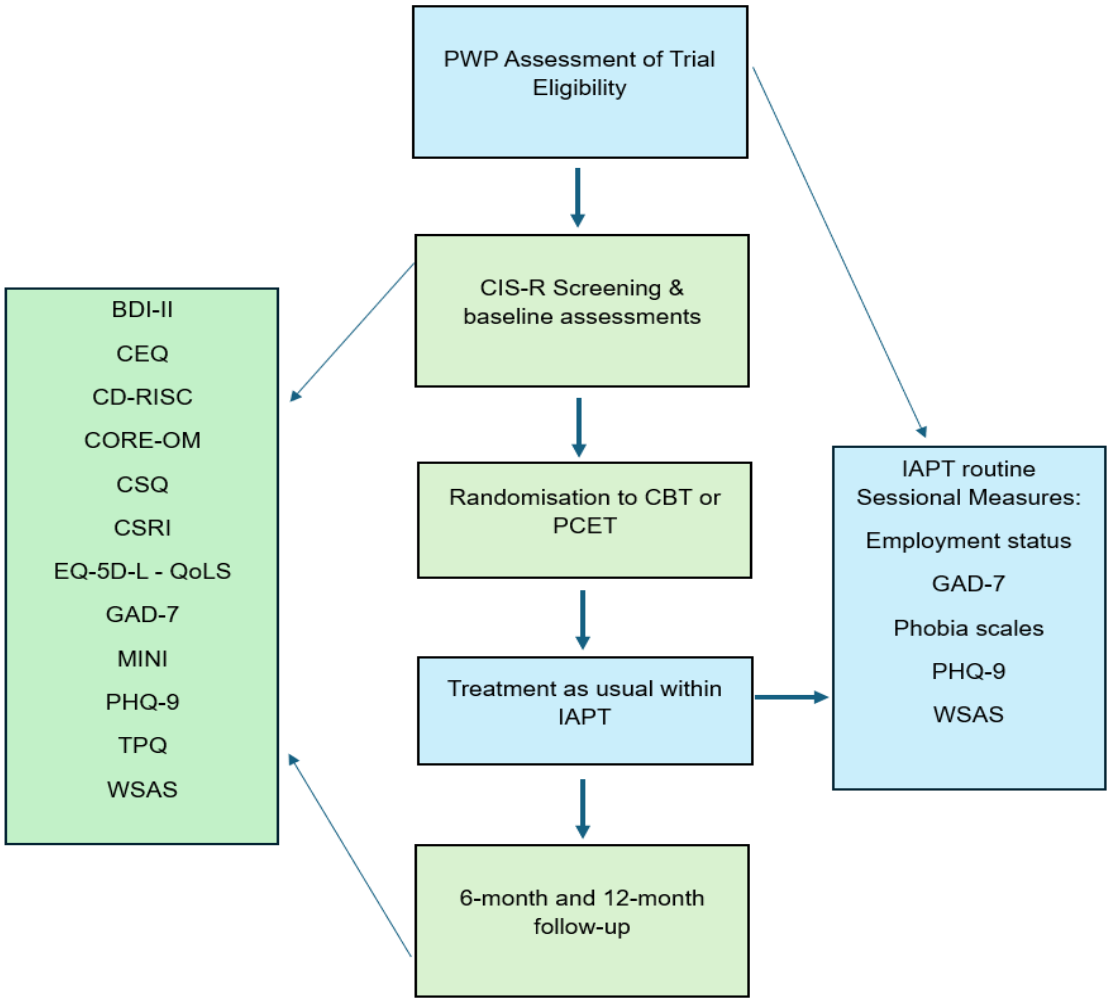
CBT is a time-limited and structured intervention exploring the relationship between thoughts, emotions and behaviour (Fenn & Byrne, 2013). CBT aims to reduce an individual's level of distress by helping one to develop healthier cognitions and behaviours, by eliminating safety behaviours and modifying faulty beliefs, (Nakao et al., 2021). PCET (sometimes known as Counselling for depression) utilises the competencies of a humanistic approach and emotional focused therapies (Drewitt, et al., 2018).

Trial Measures

As part of the PRaCTICED trial, multiple outcome measures (baseline and sessional) were collected and reported in Figure 1. However, not all measures were used in the present study. Measures used within the current study are described in depth below and can be found in Appendix C.

Figure 1

Flow Diagram Presenting Outcome Measures administered at Varying Stages of the PRaCTICED Trial.



Abbreviations: BDI-II Beck Depression Inventory II (Beck et al., 1996); CBT = Cognitive Behavioural Therapy; CEQ = Credibility & Expectancy Questionnaire (Devilly & Borkovec, 2000); CD-RISC = Connor-Davidson Resilience Scale (Connor & Davidson, 2003); CISR = Clinical Interview Schedule-Revised (Lewis et al., 1992); CSQ = Client Satisfaction Questionnaire (Attkisson & Zwick (1982); CSRI = Client socio-demographic and service receipt inventory (Chisholm et al., 2000) CORE-OM = Clinical Outcomes in Routine Evaluation-Outcome Measure (Evans et al., 2002); EQ-5D-L (Herdman et al., 2011); GAD-7 = Generalised Anxiety Disorder Measure (Spitzer et al., 2006); IAPT = Improving Access to Psychological Therapies; MINI = The Mini-International Neuropsychiatric Interview (Sheehan et al., 1998); PHQ-9 = Patient Health Questionnaire-9 (Kroenke et al., 2001); PCET = Person-centred experiential therapy; PWP = Psychological Well-being Practitioner; TPQ = Treatment Preference Questionnaire (Leykin et al., 2007); WSAS = Work and Social Adjustment Scale (Mundt et al., 2002).

The Current Study: PRaCTICED Subsample

Data for the current study comprised all 510 trial participants. Imputations for missing data were conducted (see section below for details), following which data sets were created for the two research questions.

Dropout Study Dataset

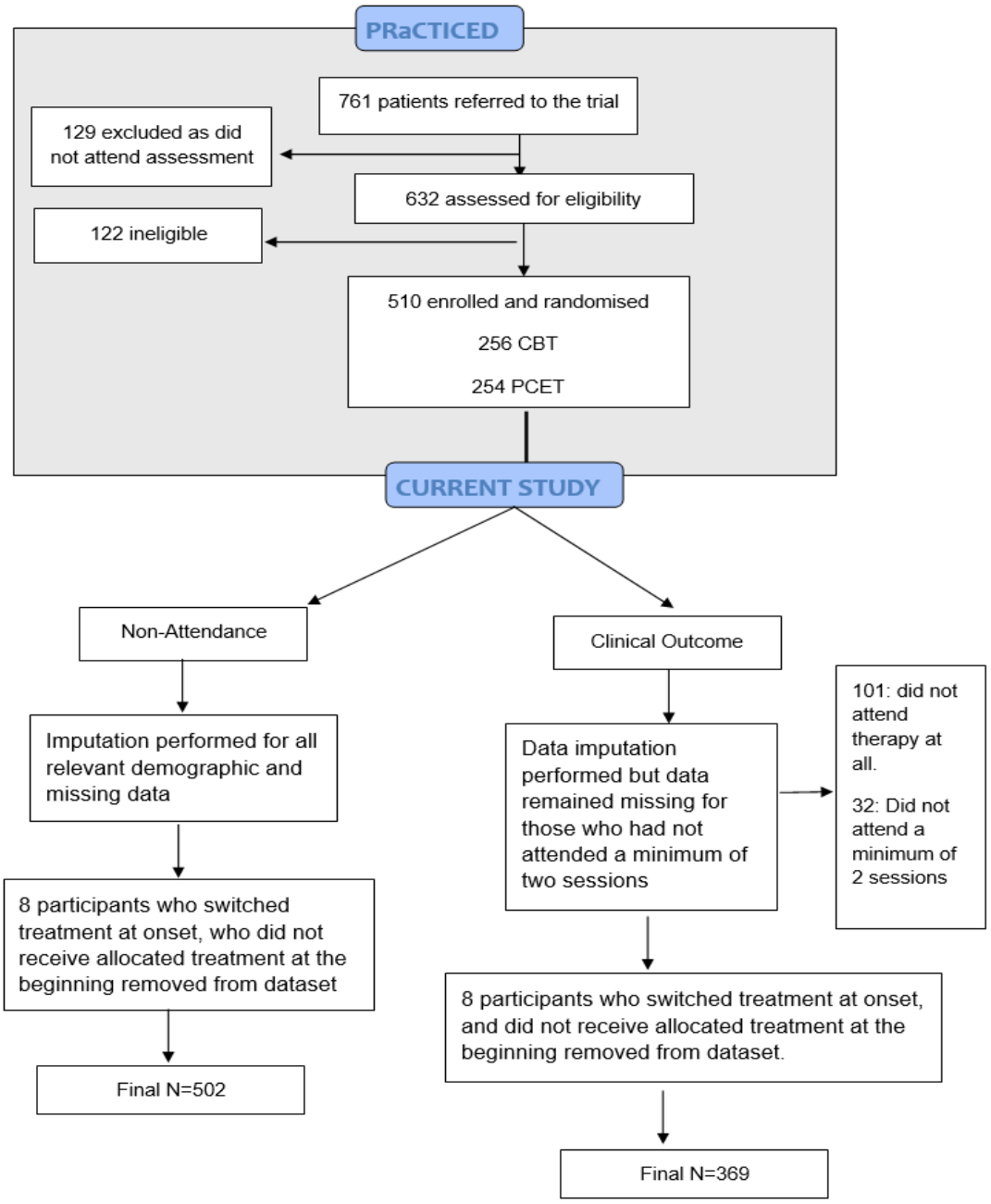
Applying the inclusion criterion to the imputed data set yielded a final sample comprising 502 participants (see Figure 2). The inclusion criterion was that clients had to have received their allocated intervention. Eight clients switched treatment at the start of therapy and were therefore removed from the analysis.

Clinical Outcomes Study Dataset

Applying the inclusion criteria, the final sample was 369 participants. The inclusion criteria were 1) those who did not attend a minimum of two therapy sessions (classified as having not received a course of treatment, as outlined by The National IAPT criteria (National Collaborating Centre for Mental Health, 2021), N = 133), and 2) participants who switched treatment at the beginning of therapy, N = 8) (see Figure 2).

Figure 2

Sub-Set Sample Eligibility Flow Chart



Study Participants

The demographic characteristics of participants are outlined in Table 1. Most participants were White-British, middle-aged, and in employment, and there were more females than males.

Table 1

Demographics and Clinical Characteristics of the Study Sample

Demographic/Clinical Category	***N*=510**	***N*=502**	**Missing N for 510 (502)**
Demographics			
Gender n, (%)			
Male	217 (42.5)	215 (42.8)	0 (0)
Female	293 (57.5)	287 (57.2)	
Age, M (SD)	38.16 (36.50)	37.95 (13.03)	0 (0)
Ethnicity n, (%)			
White	451 (88.4)	443 (88.2)	26 (26)
Other	33 (6.5)	33 (6.6)	
Employment Status, n (%)			
Employed	266 (52.2)	263 (52.4)	72 (72)
In education	34 (6.7)	34 (6.8)	
Retired	13 (2.5)	13 (2.6)	
Unemployed	124 (24.3)	119 (23.7)	
Unpaid Voluntary Work	1 (0.2)	1 (0.2)	
IMD, M (SD)	2.86 (1.58)	2.86 (1.57)	1 (1)
Referral Sector, n (%)			
Southeast	58 (11.4)	57 (11.4)	0 (0)
Southwest	257 (50.4)	251 (50.0)	
West	87 (17.1)	97 (17.3)	
North	108 (21.2)	107 (21.3)	
Clinical Characteristics			
Attendance, n (%)			
Attended Step3	409 (80.2)	401 (79.9)	0 (0)
Did not Attend Step3	101 (19.8)	101 (20.1)	
PHQ-9 Screening Score, M (SD)	18.92 (4.11)	18.91 (4.11)	0 (0)
PHQ-9 First Session (Step3) Score, M (SD)	17.22 (5.04)	17.22 (5.04)	124 (124)
PHQ-9 Last Session (Step3) Score, M (SD)	10.22 (6.53)	10.24 (6.52)	152 (149)

Wait Times (days), M (SD)	132.12 (63.50)	131.35 (62.89)	101 (101)
Outcome expectancy score of allocated treatment, M (SD)	3.37 (0.81)	3.38 (0.81)	16 (16)
Treatment preference, n (%)			
Did not mind	245 (48.0)	243 (48.4)	3 (3)
Received preferred treatment	148 (29.0)	147 (29.3)	
Did not receive preferred treatment	114 (22.4)	109 (21.7)	
Number of Step3 sessions, M (SD)	7.58 (6.71)	7.54 (6.72)	0 (0)
Allocated treatment, n (%)			
CBT	256 (50.2)	256 (50.2)	0 (0)
PCET	254 (49.8)	254 (49.8)	

Measures

All procedures and measures used in the original trial are outlined in Figure 1. The relevant study measures are described below.

Patient Health Questionnaire 9 (PHQ-9): The PHQ-9 (Kroenke et al., 2001). The PHQ-9 is a 9-item self-report, 4-point Likert Scale measure, used across many health settings to identify depression in clients (Kroenke & Spitzer, 2002). The PHQ-9 corresponds to the DSM-5 criteria for depression. Clients are required to rate how frequently they have the symptoms of depression over a 2-week period. Each item scores range from 0 (not at all) to 3 (nearly every day) with the overall score ranging from 0-27, with increasing scores indicating increasing severity. Scores of ≥ 10 are classified at meeting the clinical threshold. Scores ranging from 5-9 are categorised as mild, 10-14 are moderate, 15-19 are moderately severe and 20-27 are classified as severe levels of depression. The PHQ-9 has an internal reliability of Cronbach's α = 0.89, and a test-retest reliability of kappa of 0.84 (Kroenke et al., 2001). The PHQ-

9 was administered at screening, at the start of each therapy session, and at the end of treatment.

***Credibility and Expectancy Questionnaire*:** The Credibility and Expectancy Questionnaire (CEQ; Devilly & Borkovec, 2000), is a self-report measure comprising six items rated on a 1-9 scale or 0-100% scale dependent upon the item. However, as part of the PRaCTICED trial, the items were adapted to use 1-5 scale or 0-100% scale dependent upon the question. The initial three items of the CEQ load onto the credibility factor and the final three items load onto the expectancy factor. For this study, only the scores from the expectancy sub-scale were considered. The measure was completed at screening, prior to randomisation and treatment, with participants rating their outcome expectancy for both treatment options. For the current study, only the expectancy score for the allocated treatment was considered when studying its impact on non-attendance and clinical outcomes. All three item scores were utilised to obtain an overall outcome expectancy score of the treatment given. Items presented as percentages were converted into a 1-5 scale using predefined bands (for example, 0-20% = 1, up to 81-100% = 5), while items already rated on a 1-5 scale, were retained in their original form. The three scores from each item were summed and divided by three, to compute a single mean outcome expectancy score. The CEQ has demonstrated good reliability and validity with a Cronbach's $\alpha = .89$ overall, and $\alpha = .79$ and .90 for the expectancy factor.

***Treatment Preference Measure*:** Individuals who had a strong preference for PCET or CBT were excluded from the study at the point of referral. The treatment preference questionnaire (Leykin et al., 2007) was completed at screening, but this was not used to exclude patients. The treatment preference measure acknowledged whilst patients were aware they would be randomised; it was more to assess the

preference they may have for one of the treatments and the extent of this. Participants were instructed to circle the number under either PCET or CBT that best described their preference for the treatment of depression. If they did not have a preference, they were instructed to circle 0. For the purpose of this study, the focus was on whether the participants had/ had not received their preferred treatment and whether this impacted non-attendance and clinical outcomes.

Client Socio-Demographic and Service Receipt Inventory (CSRI): CSRI (Chisholm et al., 2000) was administered to collect socio-demographic information, such as age, gender, ethnic group, education etc.

Waiting times: Waiting times were defined as the time (days) between the Psychological Wellbeing Practitioner (PWP) stepping up the participant to step3 to the first therapy session by a step3 clinician. The association between wait times and non-attendance was not studied, due to wait time data not being available for individuals who dropped out.

It should be noted that non-attendance is defined by those who 'dropped-out' after screening, before the first session (Sweetman et al., 2023). Clinical outcome is defined as the final session score on the standardised outcome measure PHQ-9.

Retrospective Power Analyses: 'Sensitivity'

As this study used a pre-existing dataset to conduct secondary analyses, an 'a priori' power analyses were not conducted, as this type of analysis outlines the number of participants a study requires to detect a given effect size. Instead, a sensitivity power analysis was conducted which outlines what effect size the sample size is sensitive to detect. The sensitivity power calculations were conducted for each regression using G*Power 3.1.9.7 (Faul et al., 2009; see Appendix D).

The sensitivity power analysis for the multiple regression analysis indicated that with a sample of 369 participants, it would be possible to detect a minimum effect size of $f^2 = 0.045$, with the alpha set at .05, at 80% power. This implies that the study is sufficiently powered to detect small-to-medium effects in the regression model with a high likelihood of identifying true effects if present.

A further sensitivity analysis was conducted to determine whether the sample size was sufficient for a logistic regression analysis. The power analysis indicated that with a sample of 510 participants, it would be able to detect an effect corresponding to an odds ratio of 1.38 or greater, (moderate effects) with the alpha set at .05, at 80% power.

Data Analytic Strategy

The main analyses conducted were logistic regression to study non-attendance and multiple linear regression to study clinical outcomes. These analyses were conducted using IBM SPSS Statistics (Version 29; IBM Corp, 2022).

Handling Missing Data

Before conducting analyses, the missing data was imputed using the R Statistical Software (R Core Team) package MissForest algorithm (Stekhoven & Bühlmann, 2012) using the full sample of 510 participants. The process involves the use of random forest models to predict missing values, with each variable's missing values being predicted based on the observed values of other variables within the dataset (see Appendix E). MissForest, which is a non-parametric imputation method, has demonstrated its effectiveness at handling missing values in variables that have up to 30% missing information. It also has the capability to handle datasets that consist of a mix of continuous and categorical variables (Stekhoven & Buhlmann,

2012). For both research questions, analyses were conducted on both the imputed and the complete-case dataset. The complete-case dataset included all original, raw data without any imputations. Case-sensitivity analyses were run to stress-test findings, evaluate the robustness of findings, and mitigate the risk of bias in reporting.

Logistic Regression

One logistic regression was conducted on the imputed data set, whilst a second regression was conducted using the complete-case data set, as part of a sensitivity analysis. The ‘forced entry’ approach was taken whereby all exploratory variables (treatment preference and outcome expectancy), and control variables were entered into the model simultaneously to assess whether they predicted non-attendance, which was labelled as 0: attended and 1: did not attend, with the first category (0) being the reference category. Categorical predictors and control variables with two categories were labelled with 0 and 1 values (for gender: males labelled as 0, females labelled as 1; for ethnicity ‘white’ was labelled as 0 and those identifying as ‘other’ were labelled as 1. For allocated treatment, PCET was labelled as 0 and CBT was labelled as 1. The first category of each was used as reference category. Referral sector was labelled as 0: Southeast, 1: Southwest, 2: West and 3: North. Treatment preference was labelled as 0: Did not mind, 1: Received what they wanted and 2: Did not receive what they wanted. For these variables the last category was assigned as the reference category.

The assumptions for logistic regression were tested and met, for both the imputed and the complete-case dataset. Details of assumption testing for the imputed dataset are provided in Appendix F. No multicollinearity was identified as the Variance Inflation Factor VIF values were below 10 (Bowerman & O’Connell, 1990)

and tolerance values were above 0.2 (Menard, 1995) and the Pearson correlation values were below 0.7 (Field, 2018). Linearity to the logit was assessed using the Box-Tidwell procedure (Box & Tidwell, 1962) to evaluate whether there was a linear relationship between the continuous independent variables and the logit of the dependent variables. The test revealed a significant violation of linearity for the variable ‘age’ in both models, suggesting potential non-linearity. This extent of non-linearity and whether a power transformation was needed was further investigated. The following equation was used: lambda = 1 + ($\frac{b}{gamma}$): b is the estimated coefficient for the continuous predictor without the added interaction term. Gamma is the interaction between the continuous predictor variable and its natural log transformation, produced as part of the Box-Tidwell (1962) procedure. The lambda value for age for both models was calculated close to 1 which suggested there was no need for transformation, as this was indicative of the logit of the dependent variable being nearly linear. Age was retained in its original form to prevent unnecessary model over-fitting, as the non-linearity was not deemed substantial enough to impact the model’s validity.

Standardised residuals were assessed for outliers, defined by a value of ≥3.0 (Field, 2018), revealing four same cases in each model to be potential outliers. All cases were assessed, and no data entry errors were identified. Additional investigations were conducted by inspecting the Cook’s distance values, of which none were greater than 1, and leverage points which were below 0.2, Thus all values were within the acceptable range, indicating that these outliers did not significantly influence the model. Removing them led to the risk of losing important information, as they represented valid but rare values; thus, these were retained in the analysis.

Multiple Regression

One multiple regression was conducted on the imputed data set, whilst the second was ran using the complete-case data set, as part of a sensitivity analysis. The ‘forced entry’ approach was taken; the exploratory variables (wait times, outcome expectancy and treatment preference) were entered into the model to assess whether they predicted clinical outcome. Categorical predictors with two categories were labelled with 0 and 1 values, following the same approach and labelled as in the logistic regression, with the first category as the reference category. Categorical variables with three or more categories were dummy coded (Hardy, 1993). Variables with three or more categories can cause issues within multiple regression analyses, as the model can incorrectly assume an inherent order or ranking between the categories. By dummy coding, each category is treated independently. The same numeric values were assigned for variables with three or more categories as used as used in logistic regression, with the last category as the reference category.

The assumptions for multiple regression were tested for each regression analyses and met. Details of assumption testing for the imputed dataset are provided in Appendix G. Linearity was assessed using scatterplots of studentized residuals against unstandardised predicted values and partial regression plots. No multicollinearity was identified, as assessed by Pearson correlations, tolerance and VIF values. Homoscedasticity was assessed by inspecting a plot of standardised residuals against standardised predicted values and normality of residuals was assessed by P-P plots. Unusual points outliers, high leverage points and highly influential points were assessed by inspecting the case wise diagnostics, leverage

points and Cook's distance values. Independence of residuals was assessed by a Durbin-Watson statistic which was close to 2.

Moderation Analyses

Moderation analyses were conducted, only when exploratory variables were significant predictors of non-attendance and clinical outcomes. This assessed whether therapy type moderated the relationships. Firstly, continuous predictor variables were mean centred to decrease multi-collinearity issues and to aid interpretation (Iacobucci et al., 2017). Subsequently, interaction term was created by multiplying the mean-centred predictor (s) by the moderator therapy type. The created interaction terms were inserted into the initial regression models, to test for moderation effects.

Results

Non-attendance

Using the imputed data set, a logistic regression was conducted to ascertain the effects of outcome expectancy and treatment preference on non-attendance, whilst controlling for client's age, gender, ethnicity, index of deprivation, referral sector, therapy type, and PHQ-9 screening score.

Although model fit was not the primary focus, it should be noted that the logistic regression model demonstrated a statistically significant overall fit, when compared to the null model ($\chi 2(12) = 21.30$, $p = .046$, explaining 6.6% of the variance in attendance (Nagelkerke $R^2 = .066$). As seen in Table 2, the control variables, age was a significant predictor of non-attendance. For each additional year of age, the odds of non-attendance decreased by about 3% (OR = 0.97, p = .006), thus, indicating older individuals were more likely to attend in contrast to younger individuals. None of the other control variables were significant predictors.

In relation to the exploratory variables, outcome expectancy was found to be a significant predictor of non-attendance: for each one-point increase in outcome expectancy, the odds of non-attendance were 1.47 times higher (*OR* = 1.47, *p* = .011). Thus, individuals with higher outcome expectancy of treatment were more likely not to attend. Treatment preference was not a significant predictor of non-attendance (*p* > .05).

Table 2

Logistic Regression: Non-attendance (N = 502)

Variable	*B*	*SE*	*Wald*	*Df*	*P*	*OR*	*95% CI for OR*
Age	-0.03	0.01	7.60	1	.006*	0.97	[0.96, 0.99]
Gender (1)	-0.27	0.24	1.32	1	.251	0.76	[0.48, 1.21]
Ethnicity (1)	0.40	0.43	0.89	1	.345	1.50	[0.65, 3.44]
IMD	-0.15	0.08	3.23	1	.072	0.86	[0.74, 1.01]
Referral Sector: (1)	0.16	0.43	0.13	1	.715	1.17	[0.51, 2.69]
Referral Sector (2)	0.32	0.32	0.95	1	.331	1.37	[0.73, 2.59]
Referral Sector (3)	0.30	0.39	0.61	1	.436	1.35	[0.63, 2.90]
Treatment Type (1)	0.09	0.23	0.15	1	.701	1.09	[0.69, 1.73]
PHQ9 - Screening	-0.03	0.03	0.92	1	.337	0.97	[0.92, 1.03]
Treatment Pref. (1)	0.16	0.31	0.26	1	.612	1.17	[0.64, 2.16]
Treatment Pref. (2)	0.22	0.35	0.39	1	.532	1.24	[0.63, 2.44]
Outcome Expectancy	0.39	0.15	6.48	1	.011*	1.47	[1.09, 1.98]

Variable	*B*	*SE*	*Wald*	*Df*	*P*	*OR*	*95% CI for OR*
Constant	-1.10	0.91	1.48	1	.224	0.33	
Model Fit: ($\chi 2(12) = 21.30$, p= .046; .066 Pseudo R^2)							

IMD; Index of Multiple deprivation; PHQ-9; Patient Health Questionnaire. Referral Sector (1: Southeast, 2; Southwest, 3: West. North (4) = Reference category; thus, not in model); treatment preference categories: (1; did not mind/did not have preference; 2; received preferred treatment; (3) did not receive preferred treatment = reference category, thus not in model. *Note*: SPSS defaults to numbering categories starting at 1 as seen in the above output. As explained in data analysis strategy, the original category labels began at 0. .066 Pseudo R^2 statistic used was Nagelkerke R Square. *Significant at $p < 0.05$

Moderation Analysis: Non-attendance

A moderation analysis was conducted on the imputed data set to examine whether treatment type (CBT or PCET) moderated the relationship between outcome expectancy and non-attendance. As seen in Table 3, the moderation analysis revealed that the interaction term was not statistically significant (*OR* = 1.19, *p* = .571), indicating that allocated treatment did not moderate the relationship between outcome expectancy and non-attendance.

Table 3

Moderation Analysis: Non-attendance

Interaction Effect	*B*	*SE*	*Wald*	*df*	*p*	*OR*	95% CI for OR
Expectancy* Treatment Type	*0.17*	*0.30*	*0.32*	*1*	*.571*	*1.19*	[0.66, 2.13]

*Significant at $p < 0.05$

Sensitivity Analysis

A complete-case (n=459) sensitivity analysis was run to compare whether this yielded differing results (see Appendix H). The model was not statistically significant,

when compared to the null model ($\chi^2(12) = 19.64$, $p = .074$. The model explained 6.7% of the variance in attendance (Nagelkerke $R^2 = .067$). Nonetheless, the analysis revealed similar results, with each unit increase in age, decreasing the odds of not attending OR = 0.97, $p = .006$). None of the other controls were significant. Regarding the exploratory variables, outcome expectancy, was again found to be the only significant predictor of non-attendance, with each unit increase in expectancy, the odds of non-attendance were higher ($OR = 1.41$, $p = .030$). Similar to the imputed data set, the moderation analysis revealed the interaction term not to be statistically significant ($OR = 1.28$, $p = .427$), indicating that allocated treatment did not moderate the relationship between outcome expectancy and non-attendance (see Appendix I).

Multiple Regression: Clinical Outcome

The regression model was statistically significant $F(13,355) = 6.88$, $p < .001$. The model accounted for 20.1% of the variance in the post-treatment PHQ-9 score, (adjusted R^2 of 17.2%), indicating a moderate effect size. As displayed in Table 4, in relation to the control variables, the model revealed ethnicity to be a significant predictor, with those who identifying as an ethnicity 'other' than white, showing lower post-treatment PHQ-9 scores, in contrast to those identifying as white (B = -3.46, $t = -2.64$, $p = .009$). Index of multiple deprivation was also a significant control variable within the model, with individuals from less deprived areas (higher quintiles) showing reduced post-treatment PHQ-9 scores (B = -0.77, $t = -3.43$, $p < .001$). PHQ-9 first session score was also a significant control variable within the model, as individuals who had higher PHQ-9 score at the first session, tended to have higher PHQ-9 score at the last session (B = 0.45, $t = 6.94$, $p < .001$). The remaining controls were all non-significant predictors of outcome. In relation to variables of interest within the model,

wait times and treatment preference were not significant predictors of outcome. However, outcome expectancy was a significant predictor, with higher outcome expectancy associated with lower post treatment PHQ-9 scores (B = - 0.81, *t* = -1.99, *p* = .047).

Table 4

Standard Multiple Regression: Clinical Outcome: Imputed Data (N=369)

Variable	*B*	*SE*	*T*	*95% CI*	*β*	*p*
Constant	10.48	2.29	4.58	[5.98, 14.98]		< .001*
Age	-0.03	0.03	-1.21	[-0.08, 0.02]	-0.06	.226
Gender	-1.22	0.64	-1.90	[-2.47, 0.04]	-0.09	.058
Ethnicity ‘	-3.46	1.31	-2.64	[-6.03, -0.88]	-0.13	.009*
IMD-Quintile	-0.77	0.23	-3.43	[-1.22, - 0.33]	-0.19	< .001*
Referral Sector SE	-0.85	1.17	-0.73	[-3.15, 1.45]	-0.04	.467
Referral Sector SW	0.29	0.87	0.34	[-1.41, 1.99]	0.02	.737
Referral Sector W	-.031	1.06	-0.03	[-2.11, 2.05]	0.00	.977
Allocated treatment	-0.44	0.64	-0.69	[-1.69, 0.81]	-0.03	.490
PHQ-9 First Session	0.45	0.07	6.94	[0.32, 0.58]	0.34	< .001*
Wait times	0.00	0.01	-0.44	[-0.01, 0.01]	-0.02	.660
Preference: Didn’t mind	-0.53	0.82	-0.65	[-2.13, 1.08]	-0.04	.518
Preference: Received Preferred treatment	-0.47	0.92	-0.51	[-2.28, 1.35]	-0.03	.614
Outcome Expectancy	-0.81	0.41	-1.99	[-1.61, -0.01]	-0.10	.047*

Variable	*B*	*SE*	*T*	*95% CI*	*β*	*p*
F(13,355) = 6.88, *p* < .001. 17.2% (Adjusted R^2)						

IMD; Index of Multiple Deprivation; SE: South-East; SW; South-West; W: West
Referral Sector: Reference Category is North; PHQ-9; Patient Health Questionnaire,
Treatment preference: reference category is 'did not receive preferred treatment.
*Significant at p < 0.05

Moderation Analysis: Clinical Outcome

A moderation analysis was conducted on the imputed dataset to examine whether treatment type (CBT or PCET) moderated the relationship between outcome expectancy and the PHQ-9 final session score. As indicated in Table 5, the moderation analysis revealed that the interaction term was not statistically significant (B = 0.19, *t* = 0.24, *p* = .814) indicating that allocated treatment did not moderate the relationship between outcome expectancy and post-treatment score.

Table 5

Moderation Analysis Output

Interaction Effect	*B*	*SE*	*T*	*95% CI*	*β*	*p*
(Expectancy x treatment type)	0.19	0.80	0.24	[-1.38, 1.75]	0.02	.814

*Significant at p < 0.05

Sensitivity Analysis

A complete-case (*N=314*) sensitivity analysis was run to compare whether this yielded differing results (see appendix J). The regression model was statistically significant *F*(13,300) = 5.50, *p* < .001. The model accounted for 19.2% of the variance in the post-treatment PHQ-9 score, (adjusted R^2 of 15.7%), indicating a moderate effect size. Overall, running the sensitivity analysis on the raw data indicated more significant control variables in contrast to the imputed data set.

However, it should be noted that in terms of the exploratory variables, similar results were found. Higher outcome expectancy was associated with lower PHQ9 score (B = -1.09, *t* = -2.53, p = -.012), and treatment preference and wait times were not significant predictors.

Similar to the imputed data-set output, the moderation analysis revealed that the interaction term not statistically significant (*B* = -0.32, t = -0.38, p = .704) indicating that allocated treatment did not moderate the relationship between outcome expectancy and outcome (see Appendix K).

Discussion

This secondary analysis of a randomised controlled trial investigated whether pre-treatment outcome expectancy, treatment preference, and waiting times predicted pre-therapy non-attendance and clinical outcomes in clients meeting a criterion of moderate or severe depression. A secondary aim contingent on significant findings was to explore whether the type of therapy moderated the predictive influence of these predictors on non-attendance and clinical outcomes.

Results showed that pre-treatment outcome expectancy significantly predicted non-attendance before therapy began (after screening) with higher expectancy being associated with an increased risk of non-attendance, contradicting hypothesis # 1. Furthermore, no significant association was found between treatment preference and non-attendance before therapy, contrary to hypothesis # 2. The relationship between wait times and non-attendance was not investigated due to wait times data not being available for those who did not attend.

However, results did support hypothesis # 3, as pre-treatment outcome expectancy significantly predicted clinical outcomes, with higher expectancy being

associated with lower PHQ-9 scores at last session. No significant association was found between treatment preference and wait times with clinical outcomes contrary to hypotheses # 4 and # 5. Lastly, the secondary aim, which explored whether treatment type (CBT versus PCET) moderated the relationship between the significant associations, revealed no significant moderation effects.

The association between higher pre-treatment outcome expectancy and post-treatment outcome score is consistent with previous literature, suggesting that positive outcome expectations can enhance the benefits from treatment (Constantino et al., 2018; Price et al., 2013; Webb et al., 2013). Findings contradict other research which has found no association between the two factors (Jones et al., 2016; Visla et al., 2018). This finding supports the idea that therapy is most effective when clients have hope that the treatment will be beneficial for them (Frank, 1961). In the context of depression, outcome expectancy is particularly significant because individuals with depression often experience negative beliefs about their ability to improve (Beck et al., 2024).

This effect may be interpreted through the lens of expectancy theory (Vroom, 1964), which states that motivation and effort are amplified when one believes their efforts will lead to a successful outcome. In psychological therapies, such expectations may aid the therapeutic alliance, openness to change, and adherence to strategies, all of which contribute to symptom improvement. On the other hand, negative expectations may reduce motivation and adherence to strategies, reinforcing poor outcome through a self-fulfilling process (Merton, 1948). This finding indicates that factors present before treatment even starts, such as outcome expectancy, can play a vital role in clinical outcome, in isolation from therapist or treatment specific factors.

While higher pre-treatment outcome expectancy was linked to lower post-treatment depression outcome, it also correlated with a higher likelihood of non-attendance. These findings present a paradox. This was an unexpected finding which contradicts previous literature by Swift et al. (2012) and Norberg et al. (2011) who both found non-significant relationship between the two concepts. The current findings further contradict research by Bowker et al. (2022) who found clients who had higher expectation of treatment were more likely to start treatment.

One potential explanation for this unexpected finding is that clients with higher outcome expectations may experience increased anticipatory anxiety as therapy approaches. This anxiety could create internal pressure to meet their own expectations, which could induce stress. This could paradoxically lead to heightened stress about whether therapy will meet their expectations or fear that it may fall short. It could be argued to avoid the possibility of disappointment or perceived failure they may decide not to go through with therapy, engaging in what could be seen as pre-emptive disengagement. This aligns with rejection sensitivity, a trait associated with depression, whereby individuals anxiously anticipate or overreact to perceived or actual rejection or failure (Gao et al., 2017).

This behaviour can be further understood through the lens of the avoidance component of the approach and avoidance model in depression (Trew, 2011). According to the self-discrepancy theory (Higgins. 1987) and the regulatory focus theory (Higgins, 1997) individuals are motivated by their ideal self (hopes and dreams) or their ought self (duties and responsibilities). The ought self relates to avoidance motivation, where the aim is to prevent negative outcome. Thus, for clients with higher pre-treatment outcome expectancies, the fear of therapy not meeting their expectations could potentially be stronger than the desire to start

therapy. This fear can result in them dropping out of therapy (prevention focused approach) to avoid the potential risk of ‘failure’. Additionally, the heightened state of anticipatory anxiety, alongside the sadness that comes with depression are known trigger the behavioral inhibition system (BIS). The BIS system drives avoidance behaviours in response to perceived threats (Grupe & Nitschke, 2014; Trew, 2011).

Strengths & Limitations

This study comprises several strengths and limitations. One main strength is that the data comprised randomised routine data, thereby providing findings based in routine practice (i.e., high external validity) but randomised at the individual participant level (i.e., high internal validity). This makes findings more valid to real-world clinical practice and settings (Meldrum, 2000) and provide a unique research design. Furthermore, several confounder variables were accounted for, including demographics and depression severity which is known to impact how psychological therapies work (Amati et al., 2017), and thereby improving robustness. A further strength of this study is that it utilised validated and widely recognised measures, namely the CEQ to assess outcome expectancy and the PHQ-9 to assess clinical outcomes. This strengthens the methodology and enhances the validity of the findings.

Prior power analyses were not possible due to the use of secondary data. However, sensitivity power analysis indicated the sample size to assess both questions was sufficient to detect meaningful relationships. Assumptions underlying both the multiple regression and logistic regression analyses were checked, and no significant violations were found, increasing the robustness of the findings. A notable strength of this study is that missing data was imputed, ensuring the completeness of the dataset, and helping to maintain statistical power. Sensitivity analyses were

conducted on the full-case data, producing comparable results, supporting reliability of the findings, and confirming results were not driven by the handling of the missing data.

Whilst there are several strengths of this study, there are also some important limitations to consider. One limitation is that the original PRaCTICED trial measured outcome expectancy for both therapies (assigned and not assigned treatment) before allocation. The present study only considered the outcome expectancy score of treatment the participants received; it did not consider how participants felt about the alternative therapy. It may be that those who had a lower expectancy score of the treatment received compared to the alternative treatment, were more likely to not attend. This may explain the unexpected finding that higher pre-treatment outcome expectancy to be associated with higher risk of non-attendance prior to therapy beginning.

Given that individuals with a very strong preference who were against randomisation to therapy were excluded from the original PRaCTICED trial, the conclusion that treatment preference does not influence therapeutic outcomes or non-attendance, should be interpreted cautiously. The exclusion may have attenuated the effect to be weakened. Therefore, it remains important to consider patient's preferences in clinical practice, as dismissing these could result in harmful implication for the client.

Wait time data was not available, and having this data could have helped to understand the impact of wait time on non-attendance. Furthermore, longer wait times may increase anxiety and result in disappointment, specifically in those with higher outcome expectancies. This could have been a potential moderating factor between outcome expectancy and non-attendance.

A more general limitation is that the current sample in the study is predominantly white British. Thus, not representative of the wider population of individuals with common mental health difficulties reducing the generalisability of the findings. Research has found that individuals from ethnic minority groups have poorer mental health and increasingly negative experiences and outcomes in contrast to White British individuals (Bansal et al., 2022). Therefore, findings may not be applicable to other cultures and ethnicities, limiting clinical relevance. Nonetheless, it has been argued that samples derived from Western societies are often the least representative of global populations (Henrich et al., 2010). It should be noted that whilst therapists received supervision, it was less frequent than specified in the IAPT manual; this could have impacted the clinical outcomes.

Clinical Implications

This study outlines that pre-treatment outcome expectancy has a complex role to play in therapy, both improving clinical outcomes, whilst also increasing the likelihood of non-attendance prior to therapy beginning. To navigate this complexity, clinicians should assess outcome expectancy before treatment, either through verbal discussion or a standardised measure, although the latter would be the preferred option. Given the paradoxical findings, it might be important to utilise strategies to enhance outcome expectancy, such as motivational interviewing (Miller & Rollnick, 2013; Rubak et al., 2005). Statements such as “many clients feeling uncertain at first, but therapy often helps to address the concerns you are facing’, should be used. However, clinicians should ensure expectations are realistic and explain not everyone benefits from therapy (Greenberg, et al., 2006). Clinicians should gently adjust overly optimistic beliefs about therapy that could lead to high anticipatory anxiety, fear of failure and avoidance. Additionally, a clear rationale, explanation and

structure of therapy should be explained to the client, to help reduce anxiety, as well as increasing outcome expectancy. It could also be helpful for clients, specifically those with high and low outcome expectations, to be aware of what a first therapy session would entail. This could help to reduce pre-emptive anxiety, reduce the risk of non-attendance and increase expectations to an appropriate level. Services might consider making a brief check-in call to clients with high outcome expectancy, who are at risk of non-attendance. This could take place a few weeks prior to the first therapy appointment, to help ease any anxiety one is facing and to enhance the chance of attendance.

Future Directions

Future research should build on the evidence-base. The current research was unable to explore the impact of wait times on non-attendance. Thus, future research should look at the impact of wait times on non-attendance before therapy begins, in clients presenting with depression receiving CBT and PCET. There is also greater need to study predictors of PCET effectiveness as research within this area is sparse. Furthermore, future research should consider a client's outcome expectancy score of both allocated and non-allocated therapy to help better understand the role of this perspective in non-attendance, whilst exploring if wait times is a moderator factor. Finally, the effectiveness of strategies to manage expectations and reduce non-attendance should be studied.

Conclusion

The current study contributes to the understanding of how pre-treatment factors impact therapy outcomes and non-attendance in clients presenting with depression. While higher outcome expectancy was linked to lower post-therapy depression score, it also increased the risk of non-attendance. The findings outline

the importance of assessing and managing expectations carefully prior to treatment. Whilst limitations of the study are present, this study provides valuable insights for personalised care.

References

Amati, F., Banks, C., Greenfield, G., & Green, J. (2018). Predictors of outcomes for patients with common mental health disorders receiving psychological therapies in community settings: A systematic review. *Journal of Public Health (Oxford, England)*, *40*(3), 375–387. https://doi.org/10.1093/pubmed/fdx168

APA Presidential Task Force on Evidence-Based Practice (2006). Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271–285. https://doi.org/10.1037/0003-066X.61.4.271

Attkisson, C. C., & Zwick, R. (1982). The client satisfaction questionnaire. *Evaluation and Program Planning, 5*, 233–237. https://doi.org/10.1016/0149-7189(82)90074-X

Bansal, N., Karlsen, S., Sashidharan, S. P., Cohen, R., Chew-Graham, C. A., & Malpass, A. (2022). Understanding ethnic inequalities in mental healthcare in the UK: A meta-ethnography. PLoS Medicine, *19*(12), e1004139. https://doi.org/10.1371/journal.pmed.1004139

Beck, A. T., Rush, A. J., Shaw, B. F., Emery, G., DeRubeis, R. J., & Hollon, S. D. (2024). *Cognitive therapy of depression*. Guilford Publications.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. Psychological Corporation.

Beck, A., Burdett, M., & Lewis, H. (2015). The association between waiting for psychological therapy and therapy outcomes as measured by the CORE-OM.

The British Journal of Clinical Psychology, *54*(2), 233–248. https://doi.org/10.1111/bjc.12072

Bortolotti, B., Menchetti, M., Bellini, F., Montaguti, M. B., & Berardi, D. (2008). Psychological interventions for major depression in primary care: A meta-analytic review of randomized controlled trials. *General Hospital Psychiatry, 30*(4), 293–302. https://doi.org/10.1016/j.genhosppsych.2008.04.001

Bowerman, B. L., & O'Connell, R. T. (1990). *Linear statistical models: An applied approach.* (2nd ed.). Duxbury.

Bowker, H., Saxon, D., & Delgadillo, J. (2025). First impressions matter: The influence of initial assessments on psychological treatment initiation and subsequent drop-out. *Psychotherapy Research*, *35*(3), 368–378. https://doi.org/10.1080/10503307.2024.2308164

Box, G. E. P., & Tidwell, P. W. (1962). Transformation of the independent variables. *Technometrics*, *4*(4), 531–550. https://doi.org/10.1080/00401706.1962.10490038

Braybrooke, J., Ahn, H., Gallant, A., Ford, M., Bronstein, Y., Finkelstein, J., & Yee, A. (2007). The impact of surgical wait time on patient-based outcomes in posterior lumbar spinal surgery. *European Spine Journal, 16*(11), 1832-1839. https://doi.org/10.1007/s00586-007-0452-5

Chisholm, D., Knapp, M. R., Knudsen, H. C., Amaddeo, F., Gaite, L., & van Wijngaarden, B. (2000). Client socio-demographic and service receipt inventory--European version: Development of an instrument for international research. EPSILON study 5. European psychiatric services: Inputs linked to outcome domains and needs. *The British Journal of Psychiatry. Supplement*, (39), s28–s33. https://doi.org/10.1192/bjp.177.39.s28

Clark, D. M. (2011). Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: The IAPT experience. *International Review of Psychiatry*, *23*(4), 318–327. https://doi.org/10.3109/09540261.2011.606803

Connor, K. M., & Davidson, J. R. (2003). Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety*, *18*(2), 76–82. https://doi.org/10.1002/da.10113

Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. *Journal of Clinical Psychology*, *67*(2), 184–192. https://doi.org/10.1002/jclp.20754

Constantino, M. J., Vîslă, A., Coyne, A. E., & Boswell, J. F. (2018). A meta-analysis of the association between patients' early treatment outcome expectation and their posttreatment outcomes. *Psychotherapy, 55*(4), 473–485. https://doi.org/10.1037/pst0000169

Culyer, T., & Cullis, J. G. (1976). Some economics of hospital waiting lists in the NHS. *Journal of Social Policy*, *5*, 239-264.

Davis, C. H., Twohig, M. P., & Levin, M. E. (2023). Choosing ACT or CBT: A preliminary test of incorporating client preferences for depression treatment with college students. *Journal of Affective Disorders*, *325*, 413–420. https://doi.org/10.1016/j.jad.2022.12.097

Delgadillo, J., Moreea, O., & Lutz, W. (2016). Different people respond differently to therapy: A demonstration using patient profiling and risk stratification. *Behaviour Research and Therapy*, *79*, 15–22. https://doi.org/10.1016/j.brat.2016.02.003

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and*

Experimental Psychiatry, *31*(2), 73–86. https://doi.org/10.1016/s0005-7916(00)00012-4

Di Bona, L., Saxon, D., Barkham, M., Dent-Brown, K., & Parry, G. (2014). Predictors of patient non-attendance at Improving Access to Psychological Therapy services demonstration sites. *Journal of Affective Disorders*, *169*, 157–164. https://doi.org/10.1016/j.jad.2014.08.005

Drewitt, L., Pybis, J., Murphy, D., & Barkham, M. (2018). Practitioners' experiences of learning and implementing counselling for depression (CfD) in routine practice settings. *Counselling and Psychotherapy Research*, *18*(1), 3-13.

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE–OM. *The British Journal of Psychiatry, 180*(1), 51-60. https://psycnet.apa.org/doi/10.1080/09638239818328

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Fenn, K., & Byrne, M. (2013) The key principles of cognitive behavioural therapy. *InnovAiT*, *6*(9), 579-585. https://doi.org/10.1177/1755738012471029

Field, A. (2018). *Discovering statistic using IBM SPSS statistics*. Sage Publications.

Filippidou, M., Lingwood, S., & Mirza, I. (2014). Reducing non-attendance rates in a community mental health team. *BMJ Quality Improvement Reports*, *3*(1), u202228.w1114. https://doi.org/10.1136/bmjquality.u202228.w1114

Frank, J. D. (1961). *Persuasion and healing: A comparative study of psychotherapy.* Johns Hopkins University Press.

Gao, S., Assink, M., Cipriani, A., & Lin, K. (2017). Associations between rejection sensitivity and mental health outcomes: A meta-analytic review. *Clinical psychology review*, *57*, 59–74. https://doi.org/10.1016/j.cpr.2017.08.007

Garfield, S. L. (1994). Research on client variables in psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 190–228). Wiley.

Glass, C. R., Arnkoff, D. B., & Shapiro, S. J. (2001). Expectations and preferences. *Psychotherapy: Theory, Research, Practice, Training, 38*(4), 455–461. https://doi.org/10.1037/0033-3204.38.4.455

Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? *Clinical Psychology Review*, *26*(6), 657–678. https://doi.org/10.1016/j.cpr.2005.03.002

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. *Nature Reviews. Neuroscience*, *14*(7), 488–501. https://doi.org/10.1038/nrn3524

Hanevik, E., Røvik, F. M. G., Bøe, T., Knapstad, M., & Smith, O. R. F. (2023). Client predictors of therapy drop-out in a primary care setting: a prospective cohort study. *BMC Psychiatry*, *23*(1), 358. https://doi.org/10.1186/s12888-023-04878-7

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

Hardy, M. A. (1993). *Regression with dummy variables*. Sage.

Harnas, S. J., Knoop, H., Sprangers, M. A. G., & Braamse, A. M. J. (2024). Defining and operationalizing personalized psychological treatment - A systematic literature review. *Cognitive Behaviour Therapy*, *53*(5), 467–489. https://doi.org/10.1080/16506073.2024.2333345

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and Brain Sciences*, 33(2-3), 61-83. https://doi.org/10.1017/S0140525X0999152X

Herdman, M., Gudex, C., Lloyd, A., Janssen, M. F., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*, *20*, 1727-1736. https://doi.org/10.1007/s11136-011-9903-x

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*(3), 319–340. https://doi.org/10.1037/0033-295X.94.3.319

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*(12), 1280–1300. https://doi.org/10.1037/0003-066X.52.12.1280

Hirvonen, J., Tuominen, U., Seitsalo, S., Lehto, M., Paavolainen, P., Hietaniemi, K., Rissanen, P., Sintonen, H., & Blom, M. (2009). The effect of waiting time on health-related quality of life, pain, and physical function in patients awaiting primary total hip replacement: A randomized controlled trial. *Value in Health*, *12*(6), 942–947. https://doi.org/10.1111/j.1524-4733.2009.00544.x

Institute of Medicine (US) Committee on Quality of Health Care in America. (2001). *Crossing the quality chasm: a new health system for the 21st century*. National Academies Press.

Kanter, J. W., Rusch, L. C., & Brondino, M. J. (2008). Depression self-stigma: A new measure and preliminary findings. *The Journal of Nervous and Mental Disease*, *196*(9), 663–670. https://doi.org/10.1097/NMD.0b013e318183f8af

Kennedy S. H. (2008). Treating each and every depressed patient. *Journal of Psychopharmacology, (22*(7), 19–23. https://doi.org/10.1177/0269881108093270

Kramer, T. L., Owen, R. R., Wilson, C., & Thrush, C. R. (2003). Relationship between self-report and clinician-rated impairment in depressed outpatients. *Community Mental Health Journal*, *39*(4), 299–307. https://doi.org/10.1023/a:1024068007558

Kroenke K., & Spitzer R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals, 32*(9), 509–515.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lacobucci, D., Schneider, M. J., Popovich, D. L., & Bakamitsos, G. A. (2017). Mean centering, multicollinearity, and moderators in multiple regression: The reconciliation redux. *Behavior Research Methods*, *49*(1), 403–404. https://doi.org/10.3758/s13428-016-0827-9

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (5th ed., pp. 139–193). Wiley.

Lester, S., & Harris, S. M. (2007). Factors associated with first session nonattendance at a university-based family therapy clinic. *American Journal of Family Therapy, 35*(4), 363–376. https://doi.org/10.1080/01926180600814718

Lewis, G., Pelosi, A., Araya, R., & Dunn, G. (1992). Measuring psychiatric disorder in the community: A standardized assessment for use by lay interviewers. *Psychological Medicine, 22*(2), 465-486. https://doi.org/10.1017/s0033291700030415

Leykin, Y., DeRubeis, R. J., Gallop, R., Amsterdam, J. D., Shelton, R. C., & Hollon, S. D. (2007). The relation of patients' treatment preferences to outcome in a randomized clinical trial. *Behavior Therapy*, *38*(3), 209-217. https://doi.org/10.1016/j.beth.2006.08.002

Lutz, W. (2002). Patient-focused psychotherapy research and individual treatment progress as scientific groundwork for an empirically based clinical practice. *Psychotherapy Research, 12*(3), 251–272. https://doi.org/10.1093/ptr/12.3.251

Meldrum M. L. (2000). A brief history of the randomized controlled trial. From oranges and lemons to the gold standard. *Hematology/Oncology Clinics of North America*, *14*(4), 745–vii. https://doi.org/10.1016/s0889-8588(05)70309-9

Menard, S. (1995). *Applied logistic regression analysis.* Sage.

Merton, R. K. (1948). The self-fulfilling prophecy. *Antioch Review*, 8, 193–210.

Miller, W. R., & Rollnick, S. (2013). *Motivational interviewing: Helping people change*. Guilford Press.

Morton, A. (1995). The enigma of non-attendance: A study of clients who do not turn up for their first appointment. *Therapeutic Communities, 16*(2), 117–133.

Munder, T., Flückiger, C., Leichsenring, F., Abbass, A. A., Hilsenroth, M. J., Luyten, P., Rabung, S., Steinert, C., & Wampold, B. E. (2019). Is psychotherapy effective? A re-analysis of treatments for depression. *Epidemiology and*

Psychiatric Sciences, *28*(3), 268–274. https://doi.org/10.1017/S2045796018000355

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

Nakao, M., Shirotsuki, K., & Sugaya, N. (2021). Cognitive-behavioral therapy for management of mental health and stress-related disorders: Recent advances in techniques and technologies. *BioPsychoSocial Medicine*, *15*(1), 16. https://doi.org/10.1186/s13030-021-00219-w

National Collaborating Centre for Mental Health. (2021). *The improving access to psychological therapies manual.* NHS England. https://www.england.nhs.uk/wp-content/uploads/2018/06/the-nhs-talking-therapies-manual-v6.pdf

National Health Service. (2024). Monthly operational statistics. https://www.england.nhs.uk/long-read/monthly-operational-statistics-january-2024/

National Institute for Health and Care Excellence. (2012). Shared decision making. https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-guidelines/shared-decision-making

Norberg, M. M., Wetterneck, C. T., Sass, D. A., & Kanter, J. W. (2011). Development and psychometric evaluation of the Milwaukee psychotherapy expectations questionnaire. *Journal of Clinical Psychology*, *67*(6), 574–590. https://doi.org/10. 1002/jclp.20781

Norman, I., & Ryrie, I. (2018). *The art and science of mental health nursing: Principles and practice.* Open University Press.

Nye, A., Delgadillo, J., & Barkham, M. (2023). Efficacy of personalized psychological interventions: A systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*, *91*(7), 389–397. https://doi.org/10.1037/ccp0000820

Price, M., & Anderson, P. L. (2012). Outcome expectancy as a predictor of treatment response in cognitive behavioral therapy for public speaking fears within social anxiety disorder. *Psychotherapy, (49*(2), 173–179. https://doi.org/10.1037/a0024734

Punton, G., Dodd, A. L., & McNeill, A. (2022). 'You're on the waiting list': An interpretive phenomenological analysis of young adults' experiences of waiting lists within mental health services in the UK. *PloS One*, *17*(3), https://doi.org/10.1371/journal.pone.0265542

Reichert, A., & Jacobs, R. (2018). The impact of waiting time on patient outcomes: Evidence from early intervention in psychosis services in England. *Health Economics, 27*(11), 1772-1787. https://doi.org/10.1002/hec.3800

Rubak, S., Sandbaek, A., Lauritzen, T., & Christensen, B. (2005). Motivational interviewing: a systematic review and meta-analysis. *The British Journal of General Practice*, *55*(513), 305–312.

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

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The mini-international neuropsychiatric interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59*, 22–57.

Slade M. (2017). Implementing shared decision making in routine mental health care: *World Psychiatry*, *16*(2), 146–153. https://doi.org/10.1002/wps.20412

Smith, J., Kyle, R. G., Daniel, B., & Hubbard, G. (2018). Patterns of referral and waiting times for specialist child and adolescent mental health services. *Child and Adolescent Mental Health*, *23*(1), 41–49. https://doi.org/10.1111/camh.12207

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Stekhoven, D. J., & Bühlmann, P. (2012). MissForest--non-parametric missing value imputation for mixed-type data. *Bioinformatics, 28*(1), 112–118. https://doi.org/10.1093/bioinformatics/btr597

Sweetman, J., Knapp, P., McMillan, D., Fairhurst, C., Delgadillo, J., & Hewitt, C. (2023). Risk factors for initial appointment non-attendance at Improving Access to Psychological Therapy (IAPT) services: A retrospective analysis. *Psychotherapy Research: Journal of the Society for Psychotherapy Research*, *33*(5), 535–550. https://doi.org/10.1080/10503307.2022.2140616

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

Swift, J. K., Whipple, J. L., & Sandberg, P. (2012). A prediction of initial appointment attendance and initial outcome expectations. *Psychotherapy*, *49*(4), 549–556. https://doi.org/10.1037/a0029441

Trew J. L. (2011). Exploring the roles of approach and avoidance in depression: An Integrative Model. *Clinical Psychology Review*, *31*(7), 1156–1168. https://doi.org/10.1016/j.cpr.2011.07.007

Tuominen, U., Sintonen, H., Hirvonen, J., Seitsalo, S., Paavolainen, P., Lehto, M., Hietaniemi, K., & Blom, M. (2009). The effect of waiting time on health and quality of life outcomes and costs of medication in hip replacement patients: A randomized clinical trial. *Osteoarthritis and Cartilage*, *17*(9), 1144–1150. https://doi.org/10.1016/j.joca.2009.03.014

Van Dijk, D. A., Meijer, R. M., van den Boogaard, T. M., Spijker, J., Ruhé, H. G., & Peeters, F. P. M. L. (2023). Worse off by waiting for treatment? The impact of waiting time on clinical course and treatment outcome for depression in routine care. *Journal of Affective Disorders*, *322*, 205–211. https://doi.org/10.1016/j.jad.2022.11.011

Vîslă, A., Constantino, M. J., Newkirk, K., Ogrodniczuk, J. S., & Söchting, I. (2018). The relation between outcome expectation, therapeutic alliance, and outcome among depressed patients in group cognitive-behavioral therapy. *Psychotherapy Research, 28*(3), 446–456. https://doi.org/10.1080/10503307.2016.1218089

Vogel, P. A., Hansen, B., Stiles, T. C., & Götestam, K. G. (2006). Treatment motivation, treatment expectancy, and helping alliance as predictors of outcome in cognitive behavioral treatment of OCD. *Journal of Behavior Therapy and Experimental Psychiatry*, *37*(3), 247–255. https://doi.org/10.1016/j.jbtep.2005.12.001

Von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., Vandenbroucke, J. P., & Strobe Initiative. (2014). The strengthening the

reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. International Journal of Surgery, 12(12), 1495-1499. https://doi.org/10.1016/j.ijsu.2014.07.013

Vroom, V. H. (1964). *Work and motivation*. Wiley.

Webb, C. A., Kertz, S. J., Bigda-Peyton, J. S., & Björgvinsson, T. (2013). The role of pretreatment outcome expectancies and cognitive-behavioral skills in symptom improvement in an acute psychiatric setting. *Journal of Affective Disorders*, *149*(1-3), 375–382. https://doi.org/10.1016/j.jad.2013.02.016

Whiteford, H. A., Harris, M. G., McKeon, G., Baxter, A., Pennell, C., Barendregt, J. J., & Wang, J. (2013). Estimating remission from untreated major depression: A systematic review and meta-analysis. *Psychological Medicine*, *43*(8), 1569–1585. https://doi.org/10.1017/S0033291712001717

Windle, E., Tee, H., Sabitova, A., Jovanovic, N., Priebe, S., & Carr, C. (2020). Association of patient treatment preference with drop-out and clinical outcomes in adult psychosocial mental health interventions: A systematic review and meta-analysis. *JAMA Psychiatry*, *77*(3), 294–302. https://doi.org/10.1001/jamapsychiatry.2019.3750

World Health Organization. (2023). Depressive disorder (depression). World Health Organization. https://www.who.int/news-room/fact-sheets/detail/depression [accessed 13/03/2025]

Appendices

Appendix A: STROBE Checklist

STROBE checklist

This study has been written according to the "Strengthening The Reporting of Observational Studies in Epidemiology" (STROBE) guidelines. The STROBE checklist has been completed below where applicable.

	Item No	Recommendation	Page No.
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	pp. 75 - 76
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	pp.76-77
Introduction			
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported	pp. 78-83
Objectives	3	State specific objectives, including any prespecified hypotheses	pp. 83
Methods			
Study design	4	Present key elements of study design early in the paper	p. 84
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	pp. 84-86
Participants	6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	pp. 85-86; 89-90
		(b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	pp. 92-94
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	pp. 92-94;
Bias	9	Describe any efforts to address potential sources of bias	pp. 92-99
Study size	10	Explain how the study size was arrived at	pp 85-86; 89-90
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	Pp 92-94; 95-99

Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	pp. 95-99
		(b) Describe any methods used to examine subgroups and interactions	p. 99
		(c) Explain how missing data were addressed	pp. 95-96
		(d) Cohort study—If applicable, explain how loss to follow-up was addressed	N/A
		(e) Describe any sensitivity analyses	pp. 94 -95; 101-102; 104-105
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	pp. 89-90
		(b) Give reasons for non-participation at each stage	pp 89-90
		(c) Consider use of a flow diagram	p. 90
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	pp. 90-92
		(b) Indicate number of participants with missing data for each variable of interest	N/A
		(c) Cohort study—Summarise follow-up time (eg, average and total amount)	N/A
Outcome data	15*	Cohort study—Report numbers of outcome events or summary measures over time	N/A
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	pp. 99-105
		(b) Report category boundaries when continuous variables were categorized	N/A
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	pp. 94-95; 101-102; 104-105
Discussion			
Key results	18	Summarise key results with reference to study objectives	pp. 105-106
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	pp. 108-110
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	pp. 106-110

Generalisability	21	Discuss the generalisability (external validity) of the study results	pp. 108-110
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	N/A

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

Appendix B

Ethical Approval

The University Of Sheffield.

Downloaded: 15/01/2024
Approved: 12/01/2024

[redacted]
Registration number: [redacted]
Psychology
Programme: DClinPsych

Dear [redacted]

PROJECT TITLE: The Impact of Waiting Times and Treatment Preference/Expectancies on Clinical Outcome and Attendance: A secondary analysis of the PRaCTICED data set
APPLICATION: Reference Number 058286

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 C

Measures

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

PATIENT HEALTH QUESTIONNAIRE- 9 (PHQ-9)

Over the last 2 weeks, how often have you been bothered by any of the following problems? *(Use "✔" to indicate your answer)*	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 staying 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 or 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

FOR OFFICE CODING 0 +_____+_____+_____

=Total Score: _____

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?

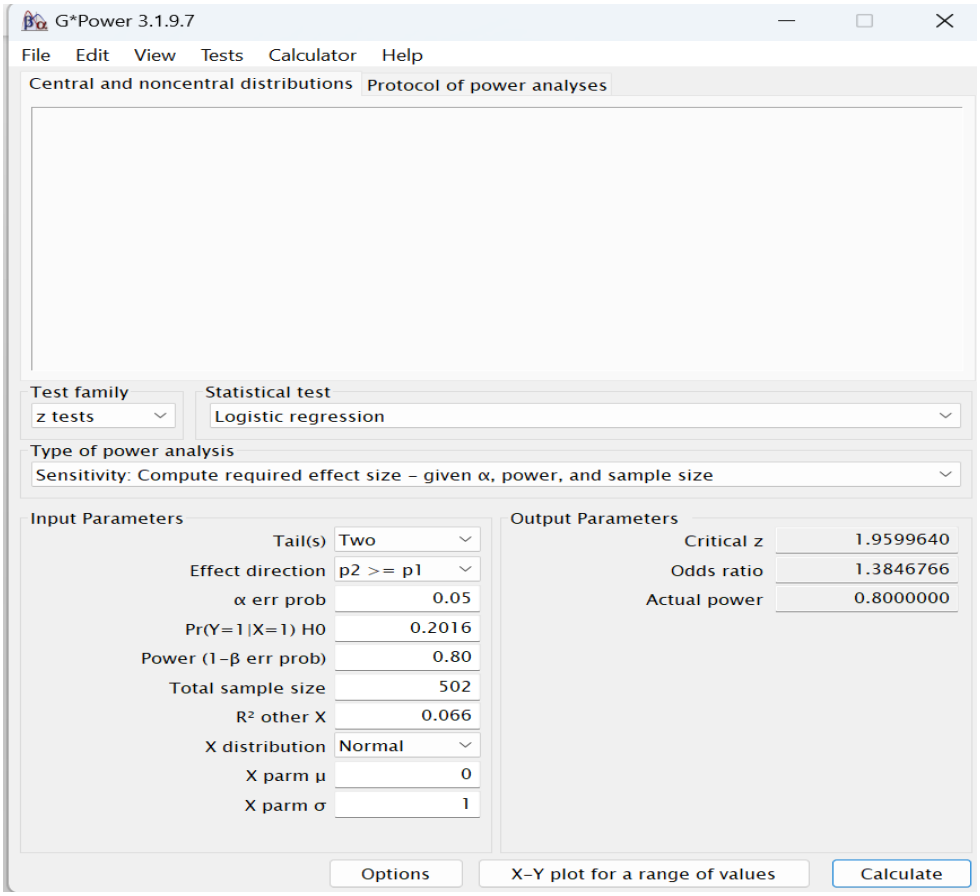
Not difficult at all	Somewhat difficult	Very difficult	Extremely difficult
☐	☐	☐	☐

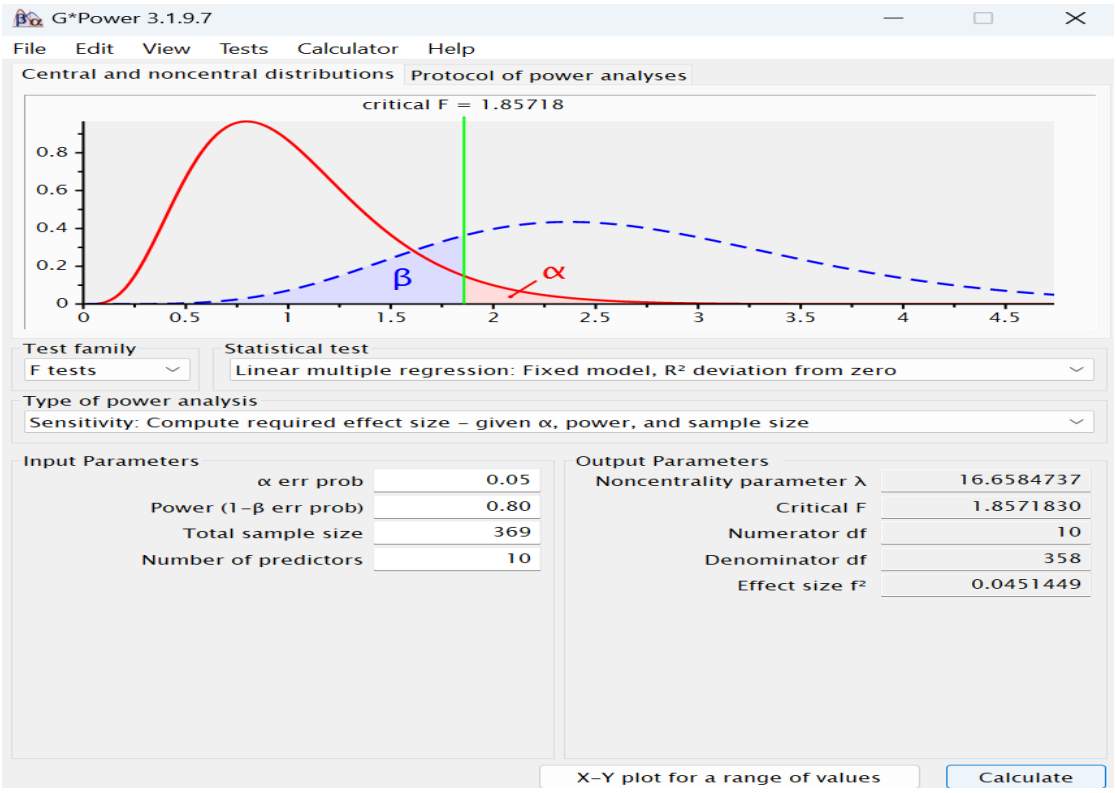
Credibility/ Expectancy Questionnaire (CEQ; Devilly & Borkovec, 2000) – Removed

Treatment Preference Measure – Removed

Appendix D

Power Analyses




G*Power 3.1.9.7
File Edit View Tests Calculator Help
Central and noncentral distributions
Protocol of power analyses
critical F = 1.85718
0.8
0.6
0.4
0.2
0
0 0.5 1 1.5 2 2.5 3 3.5 4 4.5
β
α
Test family
F tests
Statistical test
Linear multiple regression: Fixed model, R² deviation from zero
Type of power analysis
Sensitivity: Compute required effect size – given α, power, and sample size
Input Parameters
α err prob 0.05
Power (1–β err prob) 0.80
Total sample size 369
Number of predictors 10
Output Parameters
Noncentrality parameter λ 16.6584737
Critical F 1.8571830
Numerator df 10
Denominator df 358
Effect size f² 0.0451449
X–Y plot for a range of values
Calculate

Appendix E

List of Variables Used to Impute Missing Data

List of Variables Used to Impute Missing Data
Treatment Allocation
CISR
Preference Measure
Credibility Measure
Expectancy Measure
Age
Gender
IMD
Ethnicity
Employment Category
Referral Sector
PHQ-9 Measure Screening - first session; 6; Session 12; Final Session
GAD-7 Measure Screening; first session; Session 6; Session 12; Final Session
WSAS Screening; first session; Session 6; Session 12; Final Session
Phobia Measure; first session; final session
Wait times (days)
Step3 Attendance
Number of Step 3 Sessions

Appendix F

Logistic Regression: Assumption Testing: Imputed Data Set

Multicollinearity - Tolerance & VIF Values

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	.324	.142		2.278	.023		
	Age	-.004	.001	-.127	-2.801	.005	.955	1.048
	Gender	-.041	.036	-.050	-1.119	.264	.972	1.029
	Ethnicity	.071	.073	.044	.973	.331	.971	1.030
	IMDQuentile	-.019	.012	-.073	-1.610	.108	.944	1.060
	Referral Sector	-.013	.019	-.031	-.682	.496	.961	1.041
	Allocated Treatment	.013	.036	.016	.354	.724	.971	1.029
	PHQ9 at screening	-.004	.004	-.046	-1.016	.310	.966	1.036
	Whether they received preference or not	-.008	.023	-.016	-.364	.716	.954	1.049
	Expectancy Score of treatment received	.059	.023	.118	2.616	.009	.961	1.041

a. Dependent Variable: AttendanceStep3

Multicollinearity: Pearson's Correlations

Correlations

		AttendanceStep3	Age	Gender	Ethnicity	IMDQuentile	Referral Sector	Allocated Treatment	PHQ9 at screening	Whether they received preference or not	Expectancy Score of treatment received
Pearson Correlation	AttendanceStep3	1.000	-.124	-.028	.047	-.068	-.027	.023	-.022	-.026	.105
	Age	-.124	1.000	-.104	-.099	.084	.037	-.093	-.104	.011	.050
	Gender	-.028	-.104	1.000	.018	-.044	-.083	.064	.053	-.056	.028
	Ethnicity	.047	-.099	.018	1.000	-.098	.059	.007	.035	.051	-.093
	IMDQuentile	-.068	.084	-.044	-.098	1.000	-.149	.007	-.137	.025	.067
	Referral Sector	-.027	.037	-.083	.059	-.149	1.000	.015	.029	-.021	-.068
	Allocated Treatment	.023	-.093	.064	.007	.007	.015	1.000	-.030	.116	-.003
	PHQ9 at screening	-.022	-.104	.053	.035	-.137	.029	-.030	1.000	-.068	.014
	Whether they received preference or not	-.026	.011	-.056	.051	.025	-.021	.116	-.068	1.000	-.147
	Expectancy Score of treatment received	.105	.050	.028	-.093	.067	-.068	-.003	.014	-.147	1.000

Linearity to the logit - Box-Tidwell procedure

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Age	-.660	.236	7.834	1	.005	.517
	Gender(1)	-.311	.238	1.700	1	.192	.733
	Ethnicity (1)	.517	.432	1.431	1	.232	1.677
	IMDQuentile	.979	.748	1.713	1	.191	2.662
	Referral Sector			.679	3	.878	
	Referral Sector(1)	.112	.433	.067	1	.795	1.119
	Referral Sector(2)	.251	.329	.582	1	.445	1.285
	Referral Sector(3)	.267	.397	.452	1	.501	1.306
	Allocated Treatment(1)	.091	.238	.147	1	.702	1.095
	PHQ9 at screening	-.495	.716	.478	1	.489	.609
	Whether they received preference or not			.541	2	.763	
	Whether they received preference or not(1)	.218	.316	.475	1	.491	1.243
	Whether they received preference or not(2)	.230	.351	.429	1	.512	1.258
	Expectancy Score of treatment received	.506	1.840	.076	1	.783	1.659
	Age by Age_NL_Transf	.136	.051	7.251	1	.007	1.146
	IMDQuentile by IMD_NL_Transf	-.574	.375	2.346	1	.126	.563
	PHQ9Screening_NL_Transf by PHQ9 at screening	.121	.184	.433	1	.511	1.129
	Expectancy Score of treatment received by Expectancy_NL_Transf	-.037	.833	.002	1	.965	.964
	Constant	4.389	4.704	.871	1	.351	80.573

a. Variable(s) entered on step 1: Age, Gender, Ethnicity , IMDQuentile , Referral Sector, Allocated Treatment, PHQ9 at screening , Whether they received preference or not, Expectancy Score of treatment received , Age * Age_NL_Transf , IMDQuentile * IMD_NL_Transf , PHQ9Screening_NL_Transf * PHQ9 at screening , Expectancy Score of treatment received * Expectancy_NL_Transf .

Standardised residuals were assessed for outliers

Casewise List[b]

Case	Selected Status[a]	Observed AttendanceStep3	Predicted	Predicted Group	Temporary Variable		
					Resid	ZResid	SResid
24	S	D**	.134	A	.866	2.541	2.026
46	S	D**	.135	A	.865	2.535	2.020
61	S	D**	.053	A	.947	4.229	2.440
64	S	D**	.126	A	.874	2.634	2.059
105	S	D**	.130	A	.870	2.590	2.043
164	S	D**	.106	A	.894	2.904	2.139
165	S	D**	.142	A	.858	2.455	2.002
193	S	D**	.117	A	.883	2.742	2.093
214	S	D**	.130	A	.870	2.585	2.045
247	S	D**	.130	A	.870	2.585	2.039
285	S	D**	.065	A	.935	3.779	2.353
308	S	D**	.091	A	.909	3.164	2.211
334	S	D**	.131	A	.869	2.574	2.028
335	S	D**	.127	A	.873	2.617	2.051
381	S	D**	.137	A	.863	2.511	2.010
462	S	D**	.067	A	.933	3.729	2.341
470	S	D**	.131	A	.869	2.573	2.044

a. S = Selected, U = Unselected cases, and ** = Misclassified cases.

b. Cases with studentized residuals greater than 2.000 are listed.

Cook's Values

Statistics

Analog of Cook's influence statistics

N	Valid	502
	Missing	0
Mean		.0270762
Median		.0067719
Std. Deviation		.04412149
Range		.24154
Minimum		.00043
Maximum		.24197

Leverage Values

Statistics

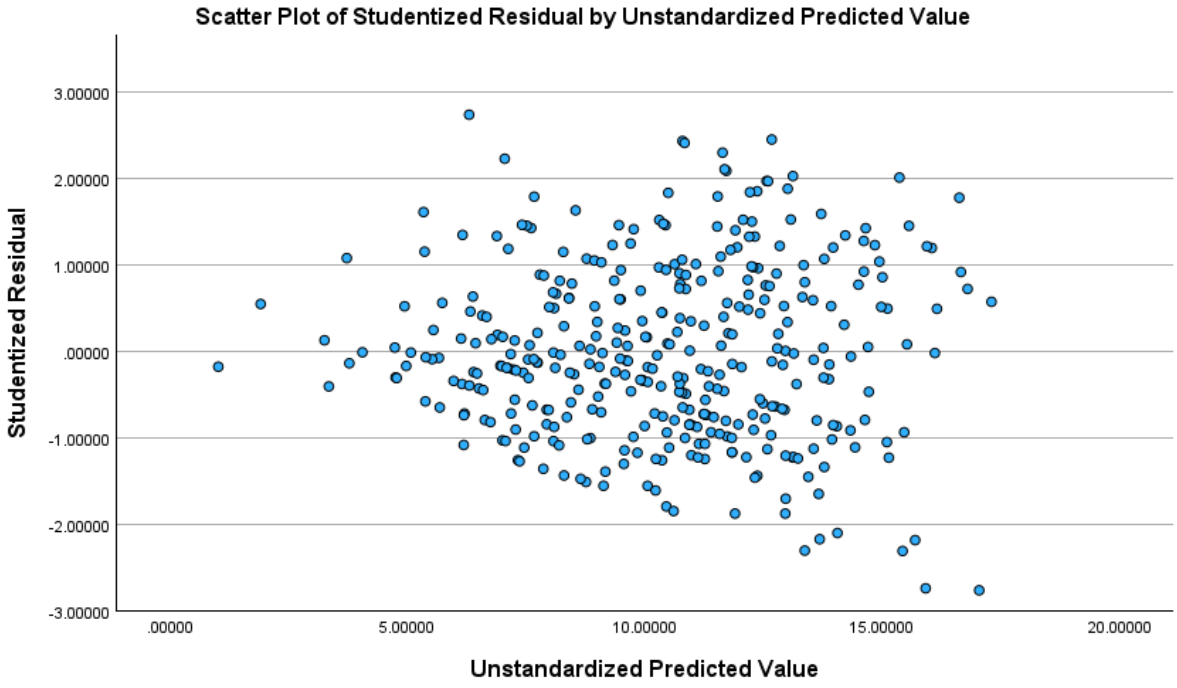
Leverage value

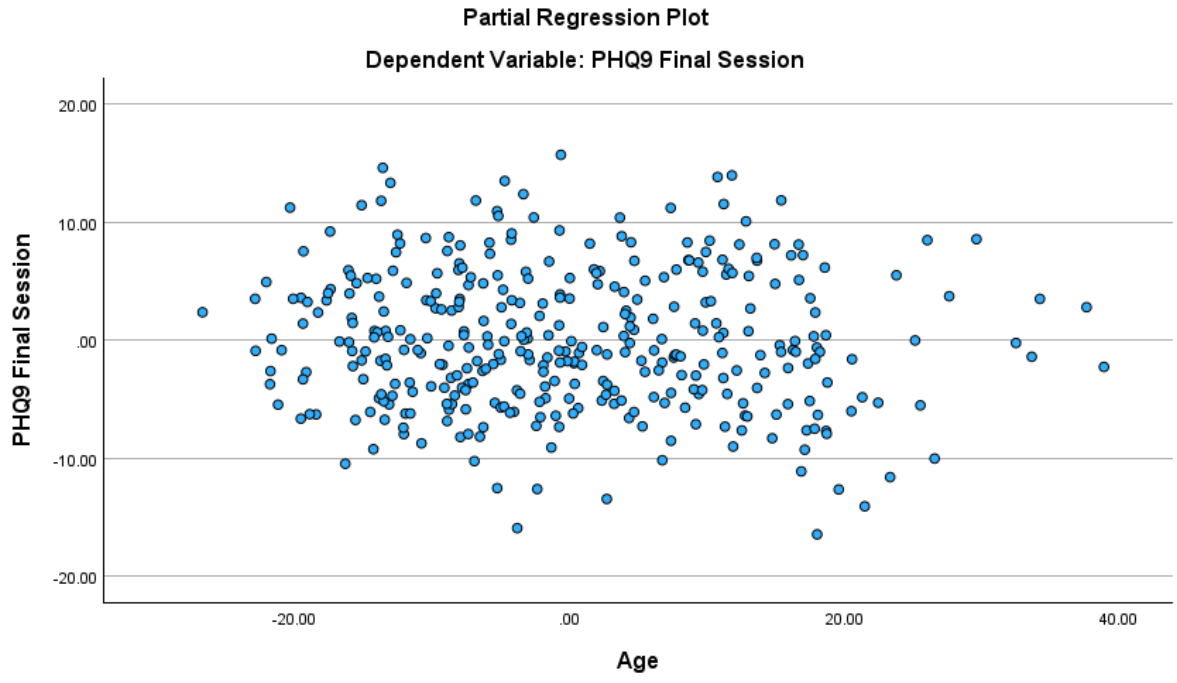
N	Valid	502
	Missing	0
Mean		.0258964
Median		.0232541
Std. Deviation		.01189988
Range		.06578
Minimum		.00956
Maximum		.07534

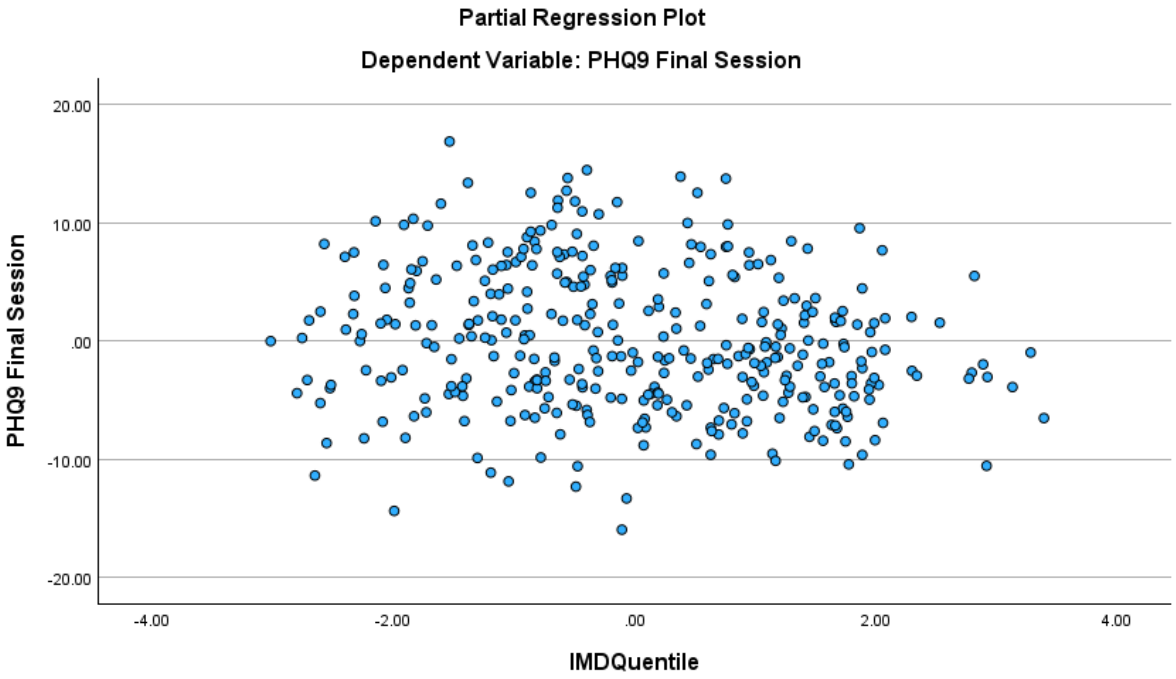
Appendix G

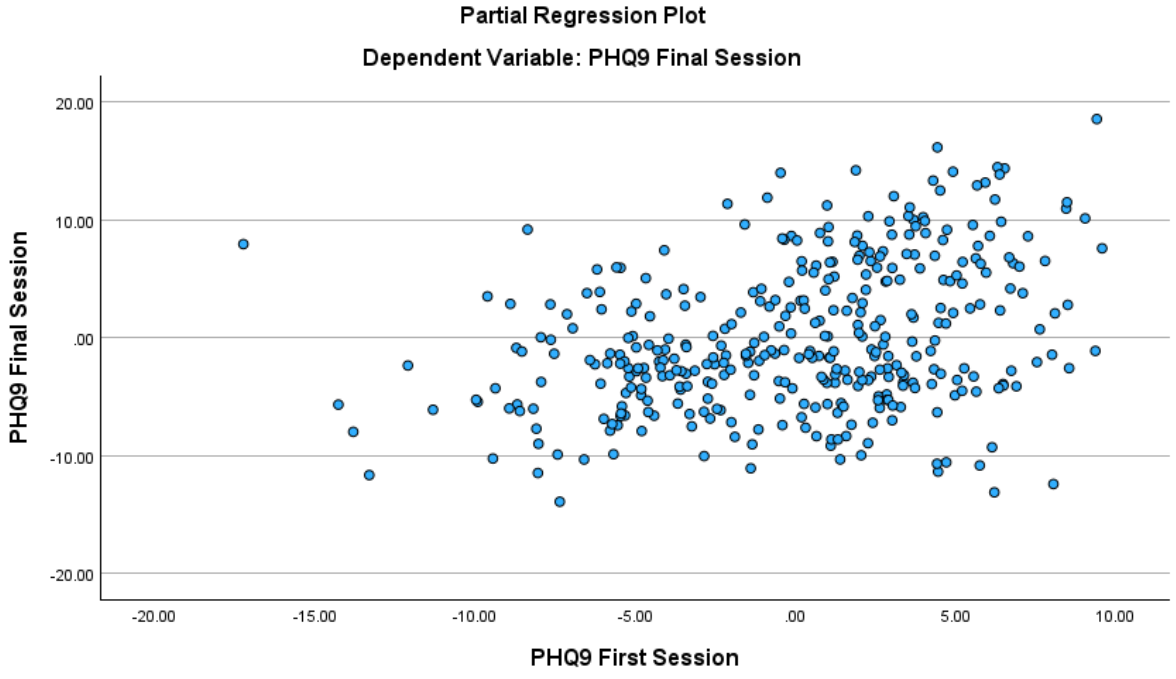
Multiple Regression: Assumption Testing: Imputed Data Set

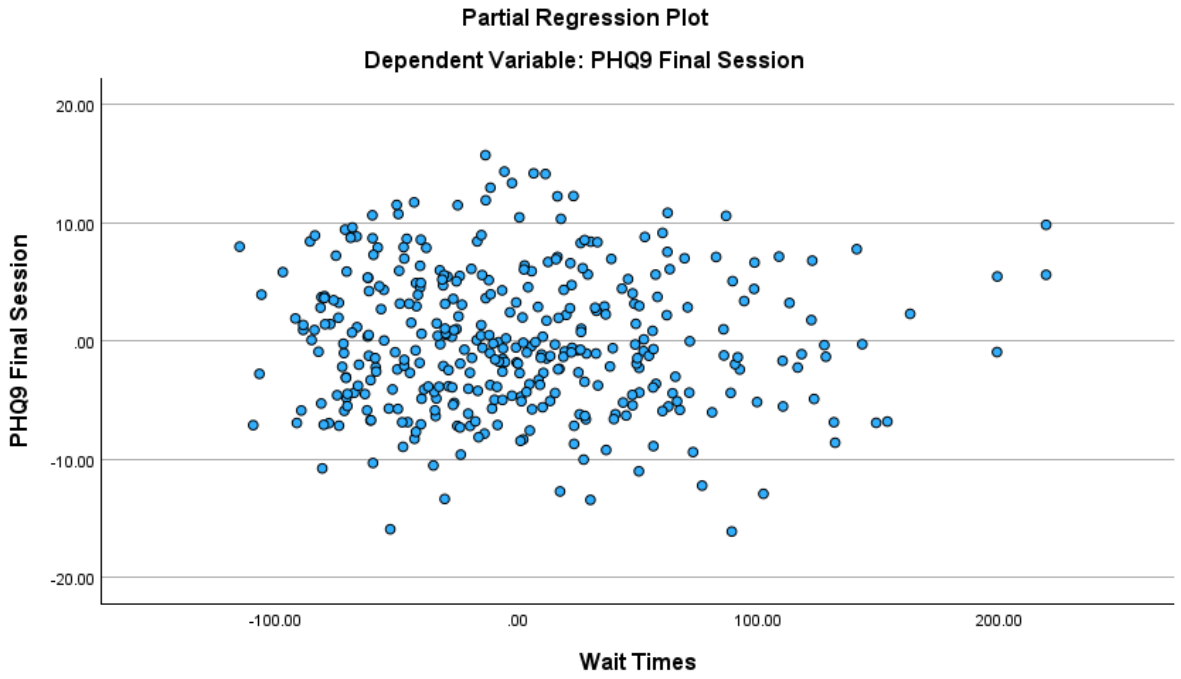
Assumption of Linearity

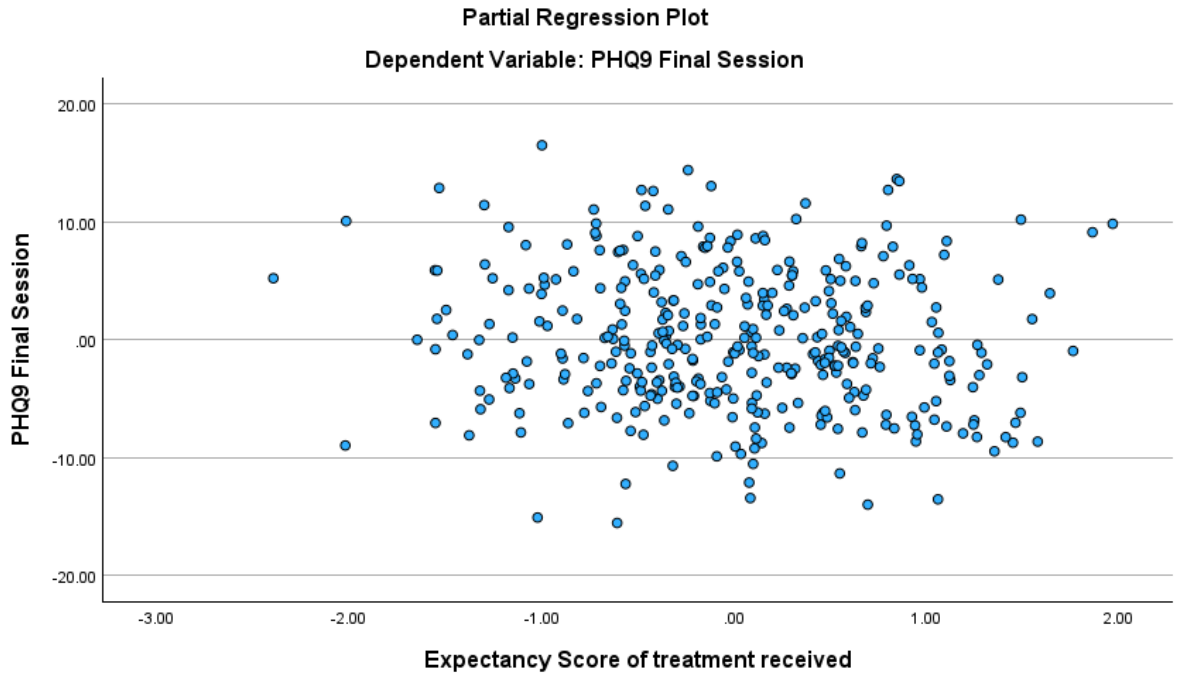





Partial Regression Plot
Dependent Variable: PHQ9 Final Session
PHQ9 Final Session
20.00
10.00
.00
-10.00
-20.00
-4.00
-2.00
.00
2.00
4.00
IMDQuentile


Partial Regression Plot
Dependent Variable: PHQ9 Final Session
PHQ9 Final Session
20.00
10.00
.00
-10.00
-20.00
-20.00
-15.00
-10.00
-5.00
.00
5.00
10.00
PHQ9 First Session


Partial Regression Plot
Dependent Variable: PHQ9 Final Session
PHQ9 Final Session
20.00
10.00
.00
-10.00
-20.00
-100.00
.00
100.00
200.00
Wait Times


Partial Regression Plot
Dependent Variable: PHQ9 Final Session
PHQ9 Final Session
20.00
10.00
.00
-10.00
-20.00
-3.00
-2.00
-1.00
.00
1.00
2.00
Expectancy Score of treatment received

Testing for Multicollinearity: Pearson's Correlation

Correlations

		PHQ9 Final Session	Age	Gender	Ethnicity	IMDQuentile	Referral_Sector=SE	Referral_Sector=SW	Referral_Sector=W	Allocated Treatment	PHQ9 First Session	Wait Times	Whether_they_received_preferred_or_not=Didn't Mind	Whether_they_received_preferred_or_not=Received what they wanted	Expectancy Score of treatment received
Pearson Correlation	PHQ9 Final Session	1.000	-.061	-.066	-.095	-.224	.011	-.025	-.027	-.042	.364	-.019	-.015	-.030	-.123
	Age	-.061	1.000	-.098	-.075	.152	-.031	.074	-.081	-.085	.023	.024	-.038	.020	.057
	Gender	-.066	-.098	1.000	-.011	-.101	.070	-.008	-.044	.045	.043	.022	.049	.043	.064
	Ethnicity	-.095	-.075	-.011	1.000	-.112	-.020	-.033	-.060	-.006	-.025	-.048	-.053	.036	-.135
	IMDQuentile	-.224	.152	-.101	-.112	1.000	-.229	.300	.143	-.042	-.162	.046	-.014	.090	.101
	Referral_Sector=SE	.011	-.031	.070	-.020	-.229	1.000	-.353	-.163	-.019	.062	-.184	.065	-.089	.070
	Referral_Sector=SW	-.025	.074	-.008	-.033	.300	-.353	1.000	-.461	-.068	-.007	.031	-.030	.092	.041
	Referral_Sector=W	-.027	-.081	-.044	-.060	.143	-.163	-.461	1.000	.016	-.043	.094	-.044	.024	-.051
	Allocated Treatment	-.042	-.085	.045	-.006	-.042	-.019	-.068	.016	1.000	-.066	.039	.030	-.162	-.022
	PHQ9 First Session	.364	.023	.043	-.025	-.162	.062	-.007	-.043	-.066	1.000	.003	.031	.012	-.018
	Wait Times	-.019	.024	.022	-.048	.046	-.184	.031	.094	.039	.003	1.000	-.082	.106	-.005
	Whether_they_received_preferred_or_not=Didn't Mind	-.015	-.038	.049	-.053	-.014	.065	-.030	-.044	.030	.031	-.082	1.000	-.622	.095
	Whether_they_received_preferred_or_not=Received what they wanted	-.030	.020	.043	.036	.090	-.089	.092	.024	-.162	.012	.106	-.622	1.000	.091
	Expectancy Score of treatment received	-.123	.057	.064	-.135	.101	.070	.041	-.051	-.022	-.018	-.005	.095	.091	1.000

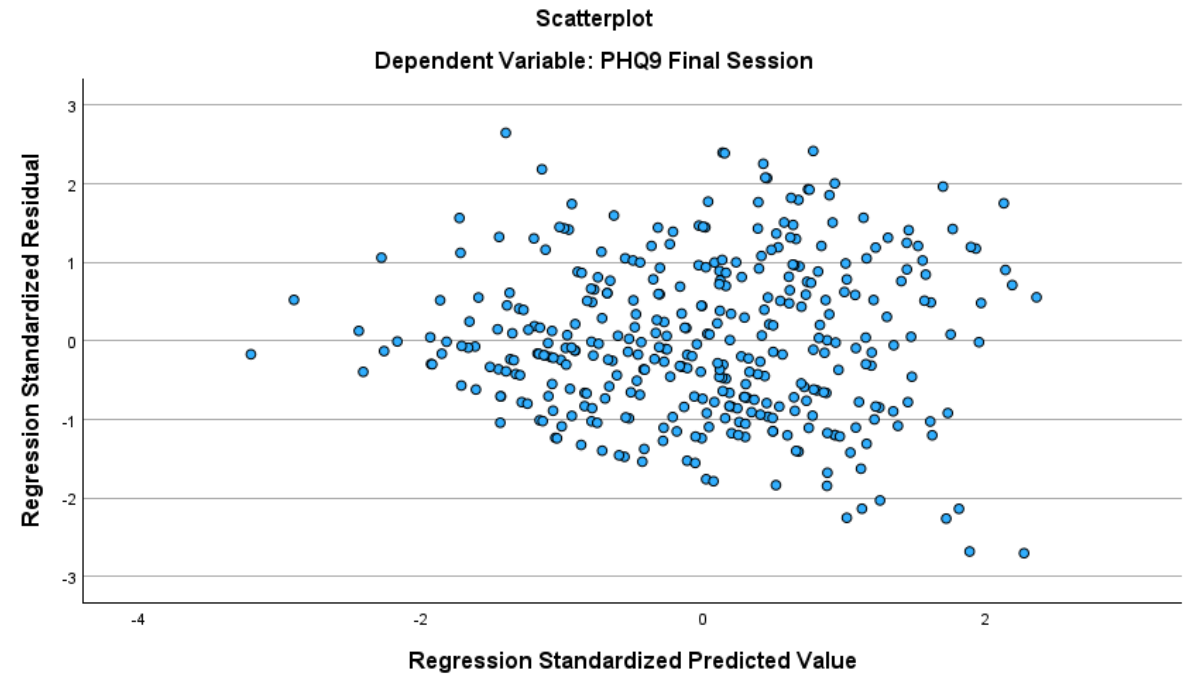
Testing for Multicollinearity: Tolerance & VIF Values

Coefficients[a]

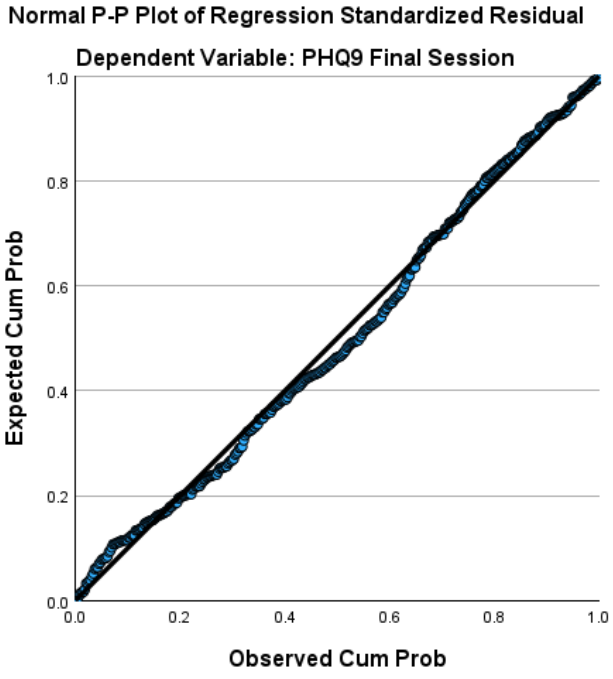
Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	10.480	2.287		4.582	<.001	5.982	14.978					
	Age	-.030	.025	-.059	-1.214	.226	-.079	.019	-.061	-.064	-.058	.940	1.064
	Gender	-1.215	.640	-.092	-1.898	.058	-2.474	.044	-.066	-.100	-.090	.959	1.043
	Ethnicity	-3.458	1.309	-.128	-2.642	.009	-6.032	-.884	-.095	-.139	-.125	.952	1.050
	IMDQuentile	-.773	.225	-.188	-3.430	<.001	-1.216	-.330	-.224	-.179	-.163	.746	1.341
	Referral_Sector=SE	-.850	1.168	-.041	-.728	.467	-3.147	1.446	.011	-.039	-.035	.708	1.412
	Referral_Sector=SW	.291	.866	.022	.336	.737	-1.412	1.994	-.025	.018	.016	.509	1.966
	Referral_Sector=W	-.031	1.058	-.002	-.029	.977	-2.111	2.050	-.027	-.002	-.001	.587	1.704
	Allocated Treatment	-.439	.636	-.034	-.691	.490	-1.690	.812	-.042	-.037	-.033	.943	1.061
	PHQ9 First Session	.450	.065	.336	6.936	<.001	.322	.578	.364	.345	.329	.957	1.045
	Wait Times	-.002	.005	-.021	-.440	.660	-.012	.008	-.019	-.023	-.021	.947	1.056
	Whether_they_received_preferred_or_not=Didn't Mind	-.527	.815	-.040	-.647	.518	-2.130	1.075	-.015	-.034	-.031	.574	1.741
	Whether_they_received_preferred_or_not=Received what they wanted	-.465	.921	-.032	-.505	.614	-2.276	1.346	-.030	-.027	-.024	.552	1.811
	Expectancy Score of treatment received	-.810	.407	-.099	-1.991	.047	-1.610	-.010	-.123	-.105	-.094	.917	1.091

a. Dependent Variable: PHQ9 Final Session

Assumption of Homoscedasticity



Normality of Residuals



Leverage Values

Statistics

Centered Leverage Value

N	Valid	401
	Missing	101
Mean		.0353802
Median		.0317207
Std. Deviation		.01414648
Range		.08911
Minimum		.01134
Maximum		.10045

Cook's Distance Values

Statistics

Cook's Distance

N	Valid	369
	Missing	133
Mean		.0029585
Median		.0013665
Std. Deviation		.00445272
Range		.03881
Minimum		.00000
Maximum		.03881

Durbin- Watson: Independence of residuals

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.449[a]	.201	.172	5.93096	2.056

a. Predictors: (Constant), Expectancy Score of treatment received , Wait Times, PHQ9 First Session, Referral_Sector=SW, Gender, Allocated Treatment, Whether_they_received_preferred_or_not=Didn't Mind, Ethnicity , Age, IMDQuentile , Referral_Sector=SE, Referral_Sector=W, Whether_they_received_preferred_or_not=Received what they wanted

b. Dependent Variable: PHQ9 Final Session

Appendix H

Logistic Regression: Predictors of Non-attendance: Raw Data (complete case; *N=459*)

Variable	*B*	*SE*	*Wald*	*Df*	*p*	*OR*	*95% CI for OR*
Age	-0.03	0.01	7.52	1	.006*	0.97	[0.95, 0.99]
Gender (1)	-0.38	0.25	2.29	1	.130	0.69	[0.42, 1.12]
Ethnicity (1)	0.36	0.45	0.63	1	.427	1.43	[0.59, 3.44]
IMD	-0.14	0.09	2.56	1	.110	0.87	[0.74, 1.03]
Referral Sector: (1)	-0.03	0.46	0.00	1	.953	0.97	[0.40, 2.39]
Referral Sector (2)	0.24	0.35	0.48	1	.487	1.27	[0.65, 2.51]
Referral Sector (3)	0.15	0.41	0.13	1	.723	1.16	[0.52, 2.60]
Treatment Type (1)	0.04	0.25	0.02	1	.884	1.04	[0.64, 1.68]
PHQ9 - Screening	-0.04	0.03	2.19	1	.139	0.96	[0.90, 1.02]
Treatment Pref. (1)	0.19	0.33	0.33	1	.568	1.21	[0.63, 2.30]
Treatment Pref. (2)	0.13	0.37	0.12	1	.731	1.13	[0.55, 2.33]
Outcome Expectancy	0.35	0.16	4.74	1	.030	1.41	[1.04, 1.93]
Constant	-0.45	0.96	0.22	1	.641	0.64	
($\chi^2(12) = 19.64$, p = .074							

Referral Sector (1: Southeast, 2; Southwest, 3: West. North (4) = Reference Category; thus, not in model); Treatment preference: (1; did not mind/did not receive preference; 2; received preferred treatment; (3) did not receive preferred treatment = reference category, thus not in model. *Note*: SPSS defaults to numbering categories

starting at 1 as seen in the above output. As explained in data analysis strategy, the original category labels began at 0 .074 Pseudo R^2 statistic used was Nagelkerke R Square. *Significant at $p < 0.05$

Appendix I

Non-Attendance: Moderation Analysis: Raw Data Set (Complete case)

Interaction Effect	*B*	*SE*	*Wald*	*df*	*p*	*OR*	95% CI for OR
Expectancy* Treatment Type	0.25	0.31	0.63	*1*	*.427*	1.28	[0.69, 2.37]

*Significant at $p < 0.05$

Appendix J

Multiple Regression Analysis: Clinical Outcome: Sensitivity Analysis (*N=314)*

Variable	*B*	*SE*	*t*	*95% CI*	*β*	*p*
Constant	12.24	2.43	5.03	[7.45, 17.03]		< .001*
Age	-0.06	0.03	-2.02	[-0.11, -0.00]	-0.11	.044*
Gender	-1.47	0.69	-2.13	[-2.83, -0.11]	-0.11	.034*
Ethnicity	-3.36	1.34	-2.51	[-6.00, -0.72]	-0.13	.013*
IMD-Quintile	-0.63	0.24	-2.61	[-1.11, -0.16]	-0.16	.009*
Referral Sector SE	-1.39	1.27	-1.10	[-3.87, 1.10]	-0.07	.273
Referral Sector SW	0.25	0.97	0.26	[-1.66, 2.16]	0.02	.799
Referral Sector W	0.07	1.14	0.06	[-2.17, 2.30]	0.00	.953
Allocated treatment	-0.91	0.68	-1.34	[-2.26, 0.43]	-0.07	.183
PHQ-9 First Session	0.42	0.07	6.05	[0.28, 0.55]	0.32	< .001*
Wait times	-0.00	0.01	-0.67	[-0.02, 0.01]	-0.04	.501
Preference - Didn't mind	0.10	0.89	0.11	[-1.65, 1.84]	0.01	.915
Preference-Received Preference	0.33	1.00	0.33	[-1.63, 2.29]	0.02	.740
Outcome expectancy	-1.09	0.43	-2.53	[-1.93, -0.24]	-0.14	.012*

$F(13,300) = 5.50$, $p < .001$
(adjusted R^2 15.7%),

IMD; Index of Multiple Deprivation; SE: South-East; SW; South-West. Referral Sector: Reference Category is North. Treatment preference: reference category is 'did not receive preferred treatment. *Significant at p < 0.05

Appendix K

Moderation Analysis Output – Sensitivity Analysis: Raw Data Set (Complete case)

Interaction Effect	*B*	*SE*	*t*	*95% CI*	*β*	*p*
(Expectancy x treatment type)	-0.32	0.83	-0.38	[-1.95, 1.32]	-0.03	.704