Do symptom change trajectories differ between Cognitive Behavioural Therapy (CBT) and Counselling? An application of growth curve modelling using a large IAPT dataset

Charles Lewis Cole

Doctorate in Clinical Psychology (DClinPsy) Thesis (Volume 1), 2021

University College London (UCL)

UCL Doctorate in Clinical Psychology

Thesis declaration form

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature: [redacted]

Name: Charles Lewis Cole

Date: 30/6/2021

Overview

Part one entails a systematic review of 45 studies that employed variations of growth curve modelling (GCM) to identify trajectory ‘classes’ of symptom change during various psychological interventions offered in primary-care and outpatient settings. Three main classes are reported across the studies: Responders, Non-responders and Deteriorators, as well as distinct subclasses of Responders based on differing rates of improvement: Rapid Responders, Delayed Responders and Unspecified Improvers. The patient characteristics associated with membership of these classes are also summarised, ranging from baseline severity to personality factors.

Part two, the empirical project, has two aims. First, it investigates the average rate of depression and anxiety symptom improvement among a large dataset of individuals receiving either CBT or Counselling. Second, the depression and anxiety trajectory classes of these individuals are defined, through application of growth mixture modelling (GMM), as: Rapid Responders, Delayed Responders, Low Severity Small Responders and Non-responders. A slower rate of depression and anxiety improvement among those who received Counselling compared to CBT was observed, as well as the finding that Counselling was associated with a lower likelihood of belonging to two classes (Rapid and Delayed Responders).

Part three is a critical appraisal of the systematic review and the empirical project. Three matters are discussed regarding the systematic review: the use of outcome feedback technologies, a tool for evaluating the reporting standards of trajectory research and the possible theoretical underpinnings of the review findings. Concerning the empirical project, reflections on the following are provided: the reasons for NICE guidance non-compliance, the process of learning GMM and the experience of working with a large dataset. Part three is concluded by an account of undertaking the thesis process during the Covid-19 pandemic.

Impact Statement

Considering their high prevalence, it is likely that everyone knows someone with a common mental disorder (CMD), such as depression or anxiety, whether that be a loved one, a colleague or themselves. Experiencing symptoms can be debilitating and can become chronic without adequate treatment, such as psychological interventions. Although evidence-based, the therapies offered in routine treatment services are not effective for everyone, with around 50% of patients continuing to exhibit clinically significant symptoms after their final session. Moreover, there is substantial variation in the rate and magnitude of symptom change over the course of treatment. Because of this, a patient's response to an intervention can be defined by their symptom 'trajectory'.

To date, many studies have explored CMD symptom change trajectories over the course of various psychological interventions, including Cognitive Behavioural Therapy (CBT), and brief psychodynamic therapy, as well as some of the patient characteristics associated with them. However, a review of these findings has not taken place nor has an investigation into the relationship between these classes and differing provisions of treatment, until now. To achieve this, a systematic review and an empirical research project was undertaken, which conferred valuable findings for patients, therapists and services, whilst inspiring future research.

The review revealed three 'classes' (trajectories of shared shape and form) of symptom change, commonly identified across the included studies. These were Responders, who could be subclassified by their rate of improvement as either rapid, delayed or unspecified, Non-Responders and Deteriorators. These classes were consistent across the various psychological interventions encompassed by the review; a finding that lends itself to the theoretical position that argues an equivalence in effect of psychological therapies. The review did, however, find that class membership was associated with various patient characteristics. Taken together, clinicians should consider these trajectories when using routine outcome monitoring (ROM) to inform treatment decisions, such as whether to continue delivering an intervention when patient progress is limited. Likewise, baseline patient characteristics could be considered when deciding on the most appropriate intervention to offer, given their associations with differing trajectory classes. This review is being prepared for submission to the peer-reviewed journal, 'Psychotherapy Research'.

The empirical project sought to determine whether the observed rates of depression and anxiety symptom improvement, differed on the basis of the intervention received (CBT or Counselling). The associations between these interventions and the trajectory classes identified were also explored. Albeit only marginally, the rate of depression and anxiety improvement was slower amongst those who received Counselling compared with CBT. In line with the review, four classes of depression and anxiety trajectories were found: Rapid Responders, Delayed Responders, Low Severity Small Responders and Non-responders. Interestingly, those who had counselling were less likely to belong to the Rapid Responder class of both depression and anxiety symptoms, as well as the Delayed Responder anxiety class. Because of this, clinicians using ROM may wish to hold in mind the potential for patients with depression to respond more slowly to Counselling than CBT and perhaps normalise the therapy journey for patients as one defined by a 'slow-burn', if they respond. Given that Counselling patients were less likely to belong to the two trajectories of anxiety symptoms with the greatest magnitude of improvements, clinicians should continue to adhere to the NICE guidelines when offering psychological intervention for anxiety disorders. The empirical paper is being prepared for submission to the peer-reviewed journal, 'Behavioural and Cognitive Psychotherapy'.

Table of Contents

Tables and Figures

Part 1: Literature Review

Part 2: Empirical Paper

Part 3: Critical Appraisal

Note. Tables 1-4 of the literature review are presented at the end of the paper due to their very large size.

Acknowledgments

It was a pleasure to complete my doctoral thesis under the guidance of my supervisors, Rob Saunders, Josh Buckman and Steve Pilling. I would like to express my gratitude for their assistance during all phases of the research process and for imparting their expertise, which enabled me to learn new skills that I will value for the rest of my career. I would also like to thank my course tutor, Josh Stott, who has supported me over the past three years when faced with professional and personal obstacles. It was also a privilege to have been taught by the UCL course team whose time and effort in delivering the training programme has been greatly appreciated.

The empirical project would not have been possible without the thousands of patients who made up the study sample. Accessing support for mental health difficulties can be a daunting, highly personal, experience and therefore, I am thankful to those who have contributed their data for the purposes of research. Likewise, I am grateful to the therapists of the North East and Central London (NCEL) IAPT services who have invested an extensive number of hours in providing care to these patients, whilst also routinely collecting their data.

Finally, I am fortunate to have had the support of my family and friends whilst undertaking the thesis, along with the camaraderie of my cohort (2018 – 2021). Above all, I would not have been able to complete my doctorate, yet alone have the opportunity to do one, if it were not for my parents, Francine and Lindsay. I am forever thankful for their care, encouragement and generosity, which has been the wind in my sail as I’ve progressed throughout my career.

Part 1: Literature Review

A Systematic Review of Common Mental Disorder Symptom Trajectories during Psychological Interventions and their Associated Patient Characteristics

Abstract

Background: Studies have used growth curve modelling (GCM) approaches to identify classes of common mental disorder (CMD) symptom change trajectories during psychological interventions, whilst the variables associated with class membership have also been explored.

Aims: The review aims to summarise the trajectory classes frequently reported by studies that employ this methodology, whilst considering whether these trajectories differ on the basis of the intervention provided, the medium of delivery and the CMD of the sample. The patient characteristics associated with the classes identified will also be summarised.

Method: A systematic search of databases was conducted to find studies that reported trajectories of symptom change during psychological interventions for various CMDs in primary-care and outpatient settings. The Latent Trajectory Studies (GRoLTS) checklist was used to assess the reporting standards of the studies included.

Results: Forty-five studies were included in the review after meeting inclusion criteria, which commonly reported the following trajectory classes: Responders, Non-responders and Deteriorators. Responders were subclassified as either Rapid Responders, Delayed Responders or Unspecified Improvers. Responders was the most frequently identified class, and in most cases, the majority of participants belonged to it. This was consistent when the trajectories were grouped by intervention type, medium of delivery and the CMD of the sample. Patient characteristics associated with these classes ranged from baseline severity to personality factors.

Conclusions: The identified classes of symptom change may be universal to all psychological interventions and various patient characteristics may be helpful in predicting membership of these. Future studies may wish to explore whether intervention type is associated with these differential classes.

Introduction

Common mental disorders (CMD) including depression, post-traumatic stress disorder (PTSD) and anxiety disorders are estimated to affect one in five adults each year and an estimated 29.2% of adults at some point throughout their lives (Steel et al., 2014). Most people with these conditions initially seek treatment in primary care, where the majority of treatment also occurs (McManus et al., 2016; Thornicroft et al., 2017). Within primary-care services, a range of psychological and psychopharmacological interventions may be offered, sometimes in combination, by adequately trained healthcare professionals (World Health Organisation [WHO], 2008). Although many types of therapy exist, those considered to be 'evidence-based' are prioritised for their supervisor efficacy, as established by gold-standard randomised controlled trials (RCT; Guidi et al., 2018). Interventions that fulfil this standard have been recommended in accordance with the CMD they are designed to treat by clinical practice guidelines such as those of The National Institute for Health and Care Excellence (NICE) in the United Kingdom.

NICE recommends that psychological therapies are considered as first-line treatments for adults with depression or anxiety (NICE, 2009). The majority of recommendations for anxiety disorders are comprised of psycho-educational groups, guided self-help, digital programmes and high-intensity therapy; all based on the principles of Cognitive Behavioural Therapy (CBT). Meanwhile, Eye Movement Desensitisation and Reprocessing (EMDR) therapy is recommended for PTSD, in addition to CBT (NICE, 2018). There are a greater number of psychological therapies recommended as first-line treatments for depression, including low- and high-intensity CBT, Interpersonal Psychotherapy (IPT), brief psychodynamic therapy, behavioural couples therapy, counselling and collaborative care (NICE, 2009). Appendix A provides a summary of the psychological interventions recommended for each CMD by NICE, taken from Clark (2018).

Despite the proven efficacy of these interventions, not all individuals who receive them have the same outcomes. Instead, a number of factors are associated with differential treatment outcomes, for instance baseline symptom severity, mental health co-morbidities and previous treatment at the individual level, and socio-economic adversity at the contextual level (Amati et al., 2018; Buckman et al., 2021; Clark et al., 2018). Therefore, researchers and clinicians alike have been encouraged to consider "what works for whom?" to ensure that patients receive treatments that are most likely to benefit them (Cohen & DeRubeis, 2018; Roth & Fonagy, 2005). In essence, this defines the 'prescriptive question', which asks whether the provision of one intervention over another is likely to benefit any given patient (Cohen & DeRubeis; Fournier et al., 2009). This differs to the 'prognostic question' that centres on how a patient's symptoms are likely to change over time, depending on whether they are given a specific treatment, a generic treatment or no treatment at all (Buckman et al., 2021). This can be explored through consideration of outcomes delineated by scores on symptom measures at the end of treatment or a study of a particular intervention. Such outcomes might include improvement, limited change, or a deterioration in symptoms.

These forms of symptom change are referred to in the literature as a 'symptom trajectory', which graphically represent the change in symptoms over a number of measurement points. These trajectories can be pooled to produce a mean rate of symptom change, known as a latent growth curve (Curran et al., 2010). The change of individuals undergoing the process of treatment can be compared to this mean, providing an estimate of their most likely outcome, which in turn can inform treatment decisions (Delgadillo et al., 2018). However, individuals receiving the same intervention can share distinct trajectories compared to others in a sample, which can be referred to a ‘classes’ of symptom change. For example, someone may respond to therapy and experience a reduction in symptoms, whereas another may not, forming the distinction between ‘responders’ and ‘non-responders’. Furthermore, some may respond very quickly compared to others who improve later in the therapy process (Saunders et al., 2019).

To explore these differential patterns of symptom change during psychological intervention, research studies have employed variations of growth curve modelling (GCM); a statistical technique that can describe both between-person and within-person change (Ram & Grimm, 2013). Commonly utilised GCM approaches that aim to identify heterogenous sub-groups based on differing symptom trajectories include growth mixture modelling (GMM), group-based trajectory modelling (GBTM) and latent class growth analysis (LGCA; van der Nest et al., 2020). Across these techniques, it is assumed that time is the within-person variable, characterised by longitudinal measurements of an outcome measure. Following the identification of distinct classes of symptom change, the association of between-subject factors with class membership can be explored via regression modelling. Together, these analyses form a ‘two-stage approach’ for researchers in highlighting sub-groups of symptom change and their respective predictors. An example of this approach relevant to psychotherapy research was conducted by Stulz and colleagues (2007), where GMM was applied to Clinical Outcomes in Routine Evaluation – Short Form (CORE-SF; Barkham et al., 2001) scores over six sessions of therapy received by 192 outpatients. Five groups of trajectories were identified: “(a) high initial impairment, (b) low initial impairment, (c) early improvement, (d) medium impairment with continuous treatment progress, or (e) medium impairment with discontinuous treatment progress” (p. 869). Furthermore, these groups were associated with differing intervention outcomes and durations, whilst age and baseline depression and anxiety severity predicted group membership.

The findings of studies that use GCM have numerous implications for clinical practice, service policy and psychological theory. First, through a greater understanding of the differential symptom trajectories observed during the early phases of interventions, researchers have been able to generate predictive models for treatment outcomes (Haas et al., 2002). In turn, these models can influence clinician treatment planning and decision-making in a manner that is beneficial for the patient, whilst reducing reliance on clinical intuition; a common reason for overestimating patient outcomes (Walfish et al., 2012). Given that GCM explores between-person differences in trajectories, the predictive strength of such an approach exceeds that of outcome systems that utilise a singular response curve comprised of patient averages, for example, the “on track” vs. “not on track” feedback system devised by Lambert and colleagues (2001). Second, identifying trajectory membership at baseline using patient pre-treatment variables, such as demographics, can allow services to develop policies surrounding care pathways. For example, if a patient is likely to be a ‘non-responder’ to psychological therapy, they may instead be supported to consider a pharmacological treatment, which have their own respective between-person trajectories (Uher at al., 2010). Finally, GCM can help inform theories regarding how and why symptoms change over time, on the basis that differences in trajectories over the duration of interventions might reflect variations in underpinning mechanisms of change (Kazdin, 2007). To date, GCM studies addressing a specific intervention type have been conducted largely on those seeking cognitive-behavioural approaches for various CMDs such as depression and some anxiety disorders, or PTSD (Joesch et al., 2013; Lutz et al., 2014; Stein et al., 2012). However, very few studies have compared differential trajectories between distinct psychological interventions in favour of conducting analyses on samples of individuals receiving various therapy types (Stulz et al., 2007).

Considering the potential benefits of using GCM approaches for improving patient care, summarising the likely symptom trajectories and the patient factors associated with them might support clinical decision making. However, only one systematic review of this nature has been conducted, focused solely on PTSD interventions, across 11 studies of civilian and veteran populations (Dewar et al., 2020). Using a narrative synthesis approach, this review found that most studies identified three distinct symptom trajectories, classified as “responders”, “non-responders” and “subclinical participants” (those who commenced therapy with symptoms that were not clinically significant), although a minority of included studies found as little as two classes and others a maximum of five. Meanwhile, 22 predictors of these trajectories were highlighted, with comorbidity of depression, anxiety or alcohol abuse serving as the strongest. Moderate predictors included combat exposure, hyper-arousal, social support and age. Unfortunately, the review was unable to explore whether the predictors of trajectories differ between different psychological interventions due to the large number included in the review. Taken together, the findings of the review provide useful insight into how symptoms of PTSD change over time for those receiving psychological interventions, which may give rise to the development of treatments that are more tailored to the individual.

The current review aims to evaluate the findings of studies of symptom change trajectories and the patient characteristics associated with these trajectories across a range of psychological interventions for adults with various CMDs, expanding on the narrow scope of Dewar and colleagues (2020). To achieve this, the following questions will be addressed:

1. What trajectories are observed across psychological interventions for CMDs?

2. Do trajectories differ between psychological interventions and their delivery medium?

3. Do symptom trajectories differ between CMD diagnoses?

4. What patient characteristics are associated with differential class membership?

Method

Search Strategy and Study Selection

The review protocol was registered with PROSPERO (2020; CRD42020212497) and is reported in accordance with the Preferred Reporting for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). The literature search took place between October 4th and 11th 2020 using a search syntax (Appendix B), comprised of key search terms, designed for each of the included databases: EMBASE, Emcare, Medline, PsycINFO, PubMed and The Cochrane Library. Search terms were combined using the AND and OR operators. The search was not restricted by year of publication. To ensure that key studies were not overlooked, an additional search was completed in Google Scholar and reference lists of articles identified through the main search were reviewed. Experts in the field were consulted for their recommendations on potentially relevant studies. In line with the Population, Intervention, Comparison and Outcome (PICO; Schardt et al., 2007) framework, the following eligibility criteria for inclusion were defined: (1) participants had a diagnosis of a CMD, or their symptoms met the clinical threshold for the disorder, indicated by validated measures. CMDs encompassed those outlined by NICE (Kendrick & Pilling, 2012): depressive disorders, generalised anxiety disorder (GAD), panic disorder, specific phobia, post-traumatic stress disorder (PTSD), obsessive compulsive disorder (OCD), social anxiety disorder (SAD), health anxiety disorder

(HAD), somatoform disorders and body dysmorphic disorder (BDD); (2) participants were adults, over the age of 18; (3) participants received a face-to-face, telephone or digital psychological intervention in a primary-care or outpatient setting, such as cognitive and behavioural therapies, brief psychodynamic psychotherapies, guided self-help interventions (GSH) and Counselling. Samples comprising of individuals receiving a combined psychological and pharmacological intervention were included; (4) CMD symptom severity across multiple time points were analysed using GCM approaches to identify multiple trajectories. If a study aimed to find multiple trajectories but only found one, its results were still included. Studies were excluded if they focused on adults with organic mental disorders, severe and enduring mental health problems, for example, psychotic disorders or neurodevelopmental disorders, such as autism spectrum disorder (ASD). Studies were also excluded if control group participants received a non-psychological intervention.

Following the search and removal of duplicates, titles and abstracts were independently screened by two reviewers (a trainee clinical psychologist and research assistant). Studies were excluded from title and abstract screening if they did not meet inclusion criteria, whilst those that did were put forward for full text screening. Once again, inclusion criteria were applied to the remaining studies after reading the full texts. If a study was excluded, the reviewer recorded the reason for this, which was later compared to their counterpart's decision. In the event of a discrepancy, the reviewers discussed the study to reach a consensus, or a third reviewer (the project supervisor) was consulted.

Data Extraction

The extraction of data was completed by two reviewers, who used Table 1 to collate data from each eligible study in line with the following: sample size, CMD(s) and recorded demographics of the sample, intervention type(s) and delivery medium, measure(s) used, GCM approach employed and time points of analysis. The trajectories identified and class membership predictors were also extracted to inform separate tables. Following this, the reviewers compared their extracted data for discrepancies, which were rectified by reviewing the study in question.

Guidelines for Reporting on Latent Trajectory Studies Checklist (GRoLTS)

There is variation in how GCM studies report identified trajectories, which can hinder the process of interpretation and critical appraisal of findings to support a systematic review. To overcome this, the current review assessed the quality of included study findings by using GRoLTS; a 16-item set of criteria to standardise the reporting of latent trajectory models (van de Schoot et al., 2017). The usability, reliability and validity of GRoLTS has been established by the authors, having been applied to the GCM analyses of 38 studies. It was found that GCM studies often overlooked six of the 13 items: (1) item 14b and c, a graph to represent the trajectories of each model investigated, alongside the means for both the final estimated trajectory model and the observed individual trajectories for each latent class; (2) item 6b, the between-class variance–covariance matrix structure; (3) item 9, the number of starting set values (4) item 16, the availability of syntax files; (5) item 3a, the missing data mechanism and (6) item 2, detailed report of time variability. Acknowledging this, each study was assigned a total score, with a maximum of 21. Because the checklist does not define thresholds of quality based on total scores, studies with low scores were not excluded, however, their findings were discussed in face of their limitations reflected by the scores they were given for individual GRoLTS items.

Data Synthesis

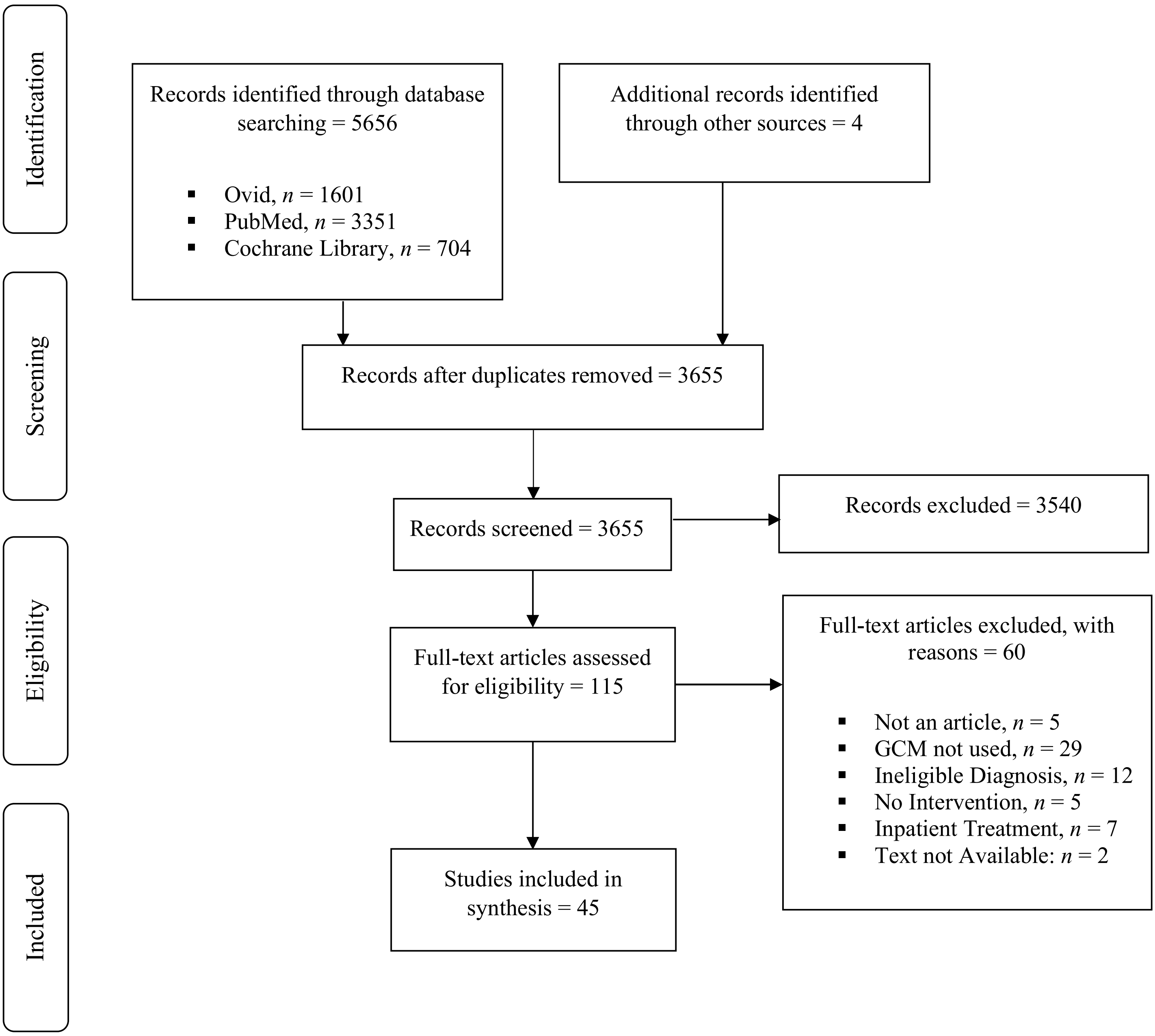
To meet the aims of the review, CMD symptom trajectories were collated on the basis of the classification assigned, i.e., the name given to the trajectory. Since the labels used to define trajectories often differ between authors, the reviewers ensured that the actual shapes and forms of trajectories were consistent through the inspection of associated graphs. If there was a mismatch between the name of a trajectory and the data it represents, the trajectory was re-classified to be consistent with those that share its shape and form. For example, some authors conceptualise trajectories of limited or partial change as ‘responders’, whilst others may view it as a ‘non-responder’. Trajectory classes that were uncommon but had a distinct shape and form with a name assigned by the authors were classified as ‘Other’. When a study explored symptom trajectories of more than one measure using separate GCM models or when a study was comprised of multiple parts, the findings were synthesised separately and distinguished by labeling the citation as ‘a’, ‘b’ or ‘c’ in the data extraction tables. Frequency counts were produced in regard to the number of trajectory classes observed across the included studies. When possible, these frequencies were presented by intervention type, CMD diagnosis and medium of delivery. Similarly, the univariate predictors of each trajectory class were collated and once again, their frequencies of association with each trajectory were recorded.

Results

The searches returned 5656 articles, whilst a further four were identified through the manual search of reference lists. No additional studies were identified via searches of Google Scholar. After removing duplicates, 3655 articles were screened based on their titles and abstracts, with 3540 removed on the basis of the exclusion criteria. This left 115 full texts, which were read in full. A further 60 articles were removed due to inclusion or exclusion criteria violations. This resulted in 45 final articles, which were then subject to data extraction (Figure 1).

Figure 1

PRISMA flowchart showing the search strategy for articles regarding symptom trajectory classes and their predictors during outpatient interventions for CMDs (Moher et al., 2009)



Included studies are presented in Table 1. These were all contemporary and published from 2007 onwards. All 45 used data from high-income countries. The sample sizes ranged from 69 to 10,854, with an average of 1149.82. The mean age of the samples ranged from 28 to 56, whilst participants were predominantly female except in six studies, which were largely comprised of male veterans with PTSD. Twenty-four studies did not report the ethnicity of their participants and for those that did, the samples were mostly of white ethnic backgrounds. Across the studies, 14 included samples of individuals with depressive disorders, participants with PTSD and anxiety disorders (mainly panic disorder with or without agoraphobia, $k = 3$) were the focus of 12 and 4 studies respectively. Twelve studies included individuals with either depression or anxiety disorders and were therefore, grouped together and referred to as 'depression and anxiety disorders'. Amongst these 12, a small number of studies included a very small proportion (< 5%) of participants with personality, substance-misuse or bipolar disorder, although their primary symptoms remained in line with the focus of the review. Three studies did not specify the CMD(s) of their samples. To clarify this, the authors were contacted for further information but no response was received. As a result, the studies were included in the present review on the basis of their primary-care settings, which provide treatment to individuals with CMDs in most cases. The overall findings of the review were considered with and without the inclusion of these studies to determine whether they affected the conclusions drawn.

The majority of studies ($k = 28$) offered an intervention that could be categorised as 'cognitive-behavioural', whereas 14 reported a variety of approaches, such as interpersonal, psychodynamic, systemic, humanistic or integrative therapy, delivered to one sample. These are referred to as 'mixed provisions' for the remainder of the review. Only two studies offered Counselling and one did not specify the intervention provided, although it was provided within a primary-care setting. These interventions were mainly offered individually and face to face (k = 31), whilst some were offered via digital programmes (k = 8) and others were delivered in a group format (k = 3). Three studies offered interventions that combined a one-to-one and group phase.

Quality Assessment – GRoLTS Scores

Overall, the quality of trajectory reporting across the included studies was acceptable, with a mean GRoLTS score of 10.50 based on the reporting standards of 43 studies (three could not be given a score due to only reporting one trajectory). The lowest GRoLTS score assigned to a study was four, in contrast to the highest of 17. No studies were excluded on the basis of a low GRoLTS score, instead, they were flagged as a source of potential limitation.

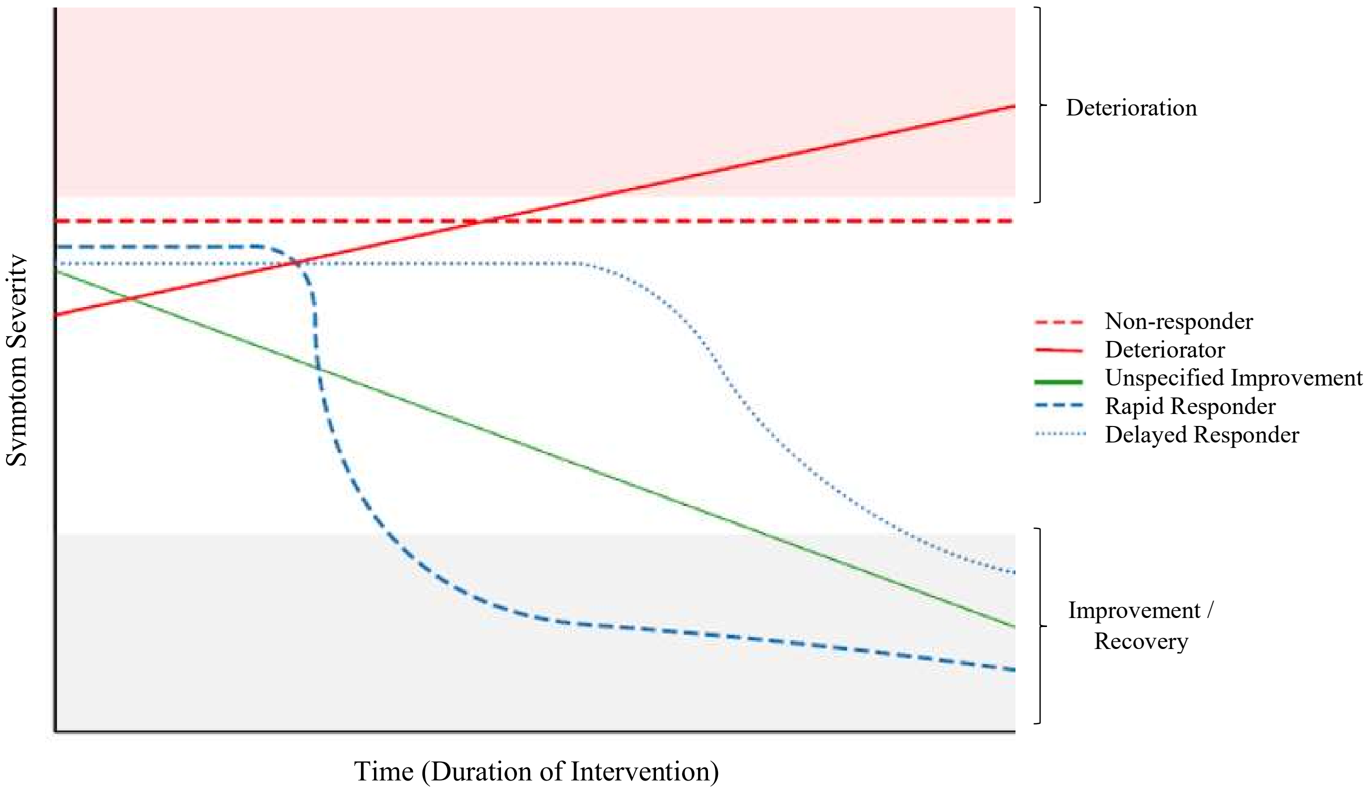
Trajectory Class Findings

In review of the included study findings, the following groups of symptom trajectory classes were identified (Table 2): three categories of 'Responders' ('Unspecified Improvement', 'Delayed or Late Response', 'Rapid or Early Response'), 'Non-responders or Limited Responders' and 'Deteriorators'. Eight classes, labeled as 'Other', did not align with these given their shape and form or were distinctive due to arising from a phased approach to therapy. These were "Improvers during group therapy", "Improvers during 1-1 therapy", "Remission Recurrence", "Early Response after Registration", "Early Response after Screening", "Early and Late Change", "Worse before Better" and "Rapid Response Remit" (Lutz et al., 2017; Moggia et al., 2020; Murphy & Smith, 2018; Owen et al., 2015; Wardenaar et al., 2014). Taken together, 163 classes were identified across the 45 studies and on average, 3.62 classes were reported per article. It is of interest to note that there is a significant correlation between the number of classes found and the sample size of each study, $r(49) = .41$, $p = .003$.

Responders were the most frequently reported trajectory class, with 105 examples reported across all studies and of these, there were 59 (36.20%) unspecified improvements, 16 (9.82%) delayed responses and 30 (18.40%) rapid responses. Regarding less favourable symptom change, there were 43 (26.80%) Non-responder classes and 7 (4.29%) Deteriorator classes. The eight 'Other' classes accounted for 4.91% of the total 163 trajectory classes identified by the review. These classes are depicted in Figure 2. Despite attempting to find multiple classes, three studies reported singular trajectories resembling average improvement and were included in the Unspecified Improvement group. As shown in Table 2, some studies identified more than one example of the same trajectory class. These were generally differentiated on the basis of baseline severity and the magnitude of symptom change.

Figure 2

A Graphical Representation of the Trajectory Classes Identified by the Review



Nine studies described themselves as exploring ‘early change’, meaning that the trajectories were only explored over the first 3-10 sessions. However, given the breadth of interventions included in the review, the utility of the term is limited because of the large degree of variation in the number of sessions offered between approaches. CBT treatments, for example, are often conducted over fewer sessions (e.g., 8-16 sessions) compared to psychodynamic psychotherapies, which commonly exceed 20 sessions. Therefore, what is considered ‘early change’ during psychodynamic therapy can encapsulate a brief CBT treatment programme in its entirety. To provide clarity on this, the timespan and measurement points of each study are stated in Table 1, ranging from three sessions to a full course of treatment with a 24-month follow-up.

Trajectory Classes by CMD. The above trajectory class frequencies were presented by CMD (Table 3). Studies comprised of individuals with either depression or an anxiety disorder (combined in one sample) accounted for the greatest proportion of the 163 total classes, followed by PTSD, and then studies of participants with depression and finally, studies of individuals with an anxiety disorder. By and large, the distribution of these CMD defined class frequencies resembled the pattern observed when considered together. Responders remained the most common trajectory class regardless of CMD, although there was some variation by sub-class, followed by Non-responders. Deteriorators and Others continued to be the least prevalent for all CMD classifications.

Trajectory Classes by Intervention Type. Table 3 also provides the class frequencies by intervention type. It was evident that CBT studies contributed the largest share of the total trajectory classes found, with mixed provisions in second and finally, studies of Counselling approaches as the smallest contributor. Once again, the observable pattern remained for each intervention category, whereby Responders formed the majority, followed by Non-responders, Others and Deteriorators. Considering that 53 trajectory classes were identified across studies involving a mixed provision of treatment, it is noteworthy that zero Deteriorators were identified and only three Delayed Responders, compared to five and 11 for CBT therapies.

Trajectory Classes by Delivery Medium. Finally, class frequencies were grouped by the medium of delivery (Table 3), referring to individual, digital, group and combined interventions (individual and group). The total 163 trajectory classes were mostly made up of those identified as part of studies that offered individual therapy, although classes were also found for studies of group, digital and even combined interventions. The same distribution of frequencies was observed on the basis of delivery medium as for CMD and intervention type; a majority of Responders, with Non-responders as the second most reported class and finally, Others and Deteriorators as the least. Because of its phased, individual and group, approach to treatment, one study reported two novel 'Other' classes. One defined by those who improved during group therapy and another characterised by improvements during one-to-one therapy only (Moggia et al., 2020).

Trajectory Class Predictors

Table 4 summarises the predictors of each of the trajectory classes found across the 45 studies. These predictors were grouped together when their associated measures explored similar constructs. As a result, there were 12 predictor groups: baseline severity, functional impairment, demographics, diagnoses, CMD symptomology, onset and course, social factors, personality factors, trauma characteristics, risk, co-morbid symptoms, physical health and treatment factors (for example, previous therapy). On most occasions, the studies used the Non-responder class as the reference group when conducting the regression analyses. The findings are reported below.

Considering baseline severity, it was apparent that higher scores on the measure used in the GCM models, indicating greater severity, predicted Non-responder status on 10 occasions and Deteriorators once. However, three studies found the inverse, whereby lower scores predicted membership of the non-responder class. Meanwhile, Rapid Responders were predicted, on the most part, by lower baseline severity, yet there were two examples where the opposite was the case. Unspecified Improvement was predicted by greater baseline scores than Non-responders, twice. Delayed Responders were not uniformly predicted by either low or high baseline scores.

There were six instances of poorer functional impairment predicting the Non-responder class and one for the Deteriorator class, whilst greater functioning generally predicted Responder membership.

Among the demographic predictors presented in Table 4, age appeared to best distinguish Responders and Non-responders, with younger participants more likely to improve over the course of intervention, although one study found that older age predicted Rapid Responder membership. On three occasions, Non-responders were predicted by unemployment, whilst belonging to a minority ethnic background and urban living predicted the class once. Finally, being married, as opposed to single, being female, and having a greater level of education appeared to predict Responder classes among the majority of studies that included these demographics.

Having a formal diagnosis of the CMD targeted by the intervention or additional co-morbid diagnoses, predicted both Responders and Non-responders on numerous occasions. Regarding the symptomology, onset and course of CMDs, three studies found that having greater levels of guilt or hyperarousal, re-experiencing and avoidance symptoms, as part of PTSD, was predictive of Non-responder class membership. Having an earlier onset of the CMD of interest was also found to predict this class.

Social Factors also appeared to have a predictive value, with Responders having greater perceived social support, social wellbeing and lower perceived burdensomeness to others. Among those with PTSD, having greater social support after the trauma distinguished Responders from Non-responders. On one occasion, Delayed Responders were found to have lower perceived social support than their Rapid Responder counterparts.

Non-responders were also predicted on the basis of several personality factors. Most notably, greater levels of neuroticism and introversion were identified as predictors of the class in two studies. Furthermore, high levels of emotion focused coping were predictive of Non-responder status, whereas Responders had lower levels, as reported by two studies. One of these studies also found that high levels of detachment-focused coping were associated with membership, as did a lower sense of mastery and greater dysfunctional attitudes as reported by one study respectively.

Among veterans receiving treatment for PTSD in four of the included studies, greater combat exposure was associated with greater likelihood of belonging to a Non-responder trajectory class. Membership of this class was also associated with a shorter time since the occurrence of the reported traumatic experience(s) for those receiving treatment for PTSD related to childhood sexual abuse.

Two studies found an association between a history of suicide attempts and greater likelihood of belonging to the Non-responder class, whilst an additional study found the same for increased levels of neglect (of self and others). Compared to Rapid Responders, Delayed Responders also had a history of more suicide attempts.

Symptoms of mental health difficulties, aside from those of the CMD of interest, were commonly reported by studies as being associated with a greater likelihood of belonging to Non-Responder trajectory classes, and conversely, lower severity of comorbid depression, insomnia, anxiety or suicidal ideation was associated with greater likelihood of belonging to a Responder class. Furthermore, Rapid Responders, when compared to Delayed or Non-Responders, had reported fewer co-morbid CMD symptoms in general.

Two studies found a relationship between having a greater number of comorbid physical health conditions and being a Non-responder, as did one for the Deteriorator class. More specifically, one study found that having fibromyalgia was associated with not responding to psychological interventions and similarly, living with HIV was found to be associated with non-response in another study. Three studies reported that those most likely to belong to delayed response trajectories had poorer physical health than those most likely to belong to rapid response trajectories.

Finally, the associations between a range of treatment related factors and trajectory class membership were explored. One study reported that past treatment, and combination treatment with an additional psychological therapy or medication were both associated with a greater likelihood of belonging to the Non-responder trajectory. A number of factors related to the way treatment was experienced or conducted were also associated with Non-responder trajectories: a poorer therapeutic alliance; a lack of group cohesion; discomfort in using the internet; and the use of written trauma accounts as part of therapy. When Delayed and Rapid Responders were compared, it was notable that those who recovered faster were more likely to have had a high- as opposed to a low-intensity intervention, in one study. Another also found that Rapid Responders were more likely to have had CBT than other forms of therapy. There were also single studies that reported that those who responded more quickly were more engaged and had greater attendance.

Discussion

Based on the findings of the 45 studies that informed the review, three main trajectory classes of CMD symptom change were found during psychological interventions in primary-care settings. These were: (1) Responders, which can be further specified as either (a) Unspecified Improvers, (b) Delayed Responders or (c) Rapid Responders, (2) Non-responders and (3) Deteriorators. Responders accounted for nearly two thirds of the total number of classes reported, whilst the remaining third was largely comprised of Non-Responders, with a minority of Deteriorators.

There were also eight trajectory classes, which were classified as 'Other' due to having a unique shape and form or the result of a novel therapy approach. Two of these, "Improvers during group therapy" and "Improvers during 1-1 therapy", refer to two groups of individuals who experienced symptom improvements dependent on the phase of treatment. In this study, all participants were offered a course of therapy that commenced with a group intervention, followed by a one-to-one provision (Moggia et al., 2020). The "Remission, Recurrence" and "Rapid Response, Remit" classes found by two independent studies, reflect the trajectories of individuals who experienced an initial improvement but later deterioration in symptoms (Murphy & Smith, 2018; Wardenaar et al., 2014). As part of a study that explored symptom trajectories over four time points during and before a web-based intervention for depression, two classes were found that reflect very early symptom improvements before and immediately after treatment sessions commenced. These are the "Early Response after Registration" and "Early Response after Screening" classes. Whilst the former reflects improvements between initial suitability screening and user registration, the latter indicates a reduction of symptoms between registration and session one (Lutz et al., 2017). The two final ‘Other’ classes labeled “Early and Late Change” and “Worse before Better” were reported by Owen and colleagues (2015). These describe individuals who experienced a plateau in symptom improvements during the middle phase of treatment and individuals who were subject to initial deterioration, which reversed by the end of the intervention.

Participants across all of the included studies were most likely to be responders to psychological intervention. This finding was consistent regardless of the CMD diagnosis or symptoms of interest, the intervention offered or the medium of its delivery. Given that the review included interventions routinely offered by primary-care or outpatient services, it was anticipated that these would have a fairly established evidence-base and therefore, be effective for the average patient. This outlook has been supported by randomised controlled trials (RCT) for a range of psychological interventions for depression and anxiety disorders, with each appearing to be similar in efficacy (Bandelow et al., 2015; Cuijpers et al., 2013; Cuijpers, 2017). However, it has been noted that in some primary-care services, therapists may offer treatments outside of the evidence-base and not in line with the protocols tested during RCTs (Clark et al., 2018). Even so, outside of the often-optimal conditions of RCTs, the effectiveness of interventions in naturalistic settings has been explored and outcomes are largely positive. Many examples of this come from evaluations of the UK’s IAPT programme, which predominantly offers CBT of high- and low-intensity, with smaller but sizeable provisions of Counselling. In these settings, a small number of patients are also offered Behavioural Couples Therapy, Collaborative Care, Interpersonal Psychotherapy (IPT), and Dynamic Interpersonal Therapy (DIT). Despite variation between regions and services, the most recent national IAPT outcomes report found 51.1% of those who referred recovered on the basis of clinical measure thresholds on both the PHQ-9 and GAD-7, whilst 67.0% showed reliable improvement (NHS Digital, 2019). Furthermore, recovery rates in terms of symptoms of just depression or anxiety disorders were higher at 55.4% and 53.4% respectively. Meanwhile, a study of IAPT delivered Counselling and CBT interventions for depression (Pybis et al., 2017) were found to be comparable in recovery and reliable improvement rates, although more recent studies have questioned this (Barkham et al., 2021; Saunders et al., 2020). These benefits of psychological intervention are not limited to the those of traditional individual formats, with many studies reporting equivalent effectiveness and efficacy findings for digital and group interventions (Burlingame et al., 2016; Fanous & Daniels, 2020; Marcelle et al., 2019; Richards et al., 2020; Stefanopolou et al., 2019). Therefore, the finding of the review that membership of the Responder class among participants receiving digital or group interventions was above and beyond that of the Non-responder or Deteriorator counterparts is unsurprising.

Among participants classified as Responders, some studies uncovered considerable heterogeneity in symptom change, resulting in two types of improvement: Rapid and Delayed Responders. Rapid Responders experienced a reduction in symptoms during the early stages of therapy, which subsequently leveled off or continued to improve. In contrast, Delayed Responders experienced a slow or negligible symptom improvement until they began to improve during later sessions. There are several reasons why someone may respond quickly to psychological intervention. First, some have attributed the rapid gains to the effect of therapy "common factors", such as the emergence of the therapeutic relationship, since the improvements occur before the introduction of therapeutic techniques. These early gains may not be present for some individuals because they are less responsive to these common factors for reasons yet known (Ilardi & Craighead, 1999). In line with this, it is important to note that the

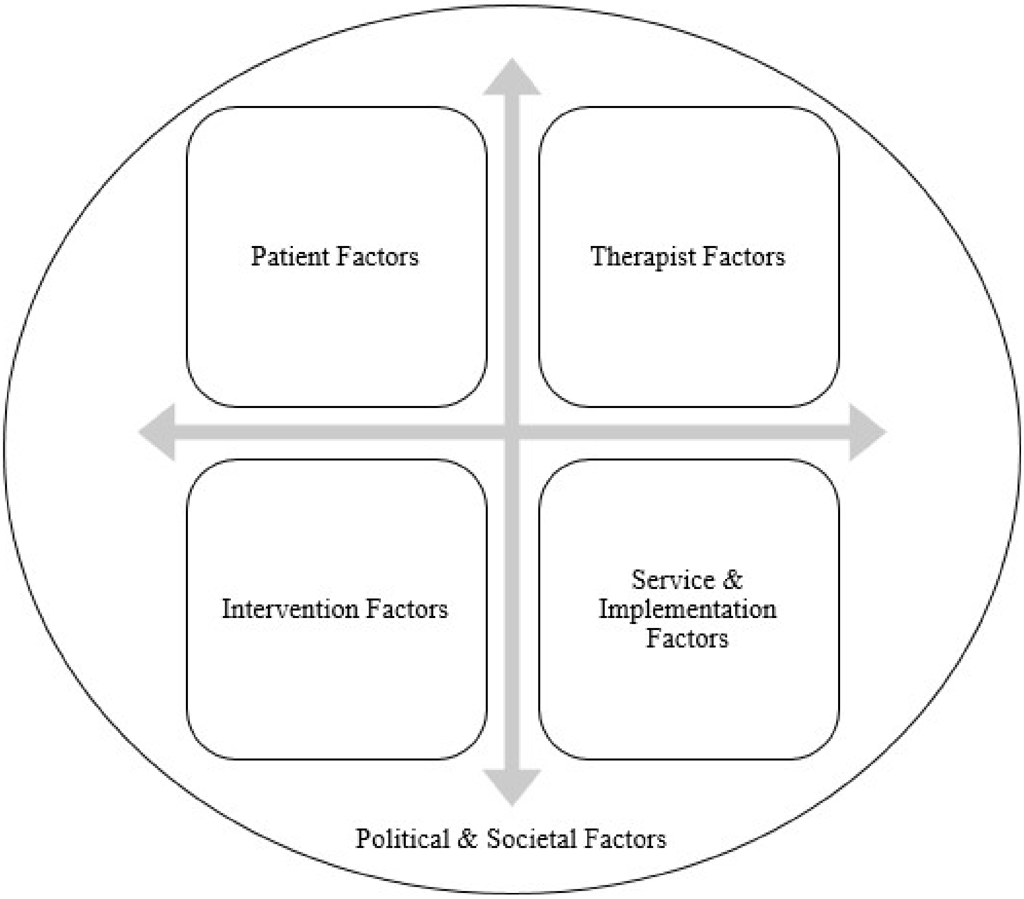
Rapid Responder class was identified across numerous studies of the review, regardless of whether they were controlled, ensuring theoretical adherence during therapy delivery, or naturalistic studies involving multiple therapy approaches. Therefore, the Rapid Responder class appears to be a cross-modality form of symptom change for many. Second, the concept of "critical sessions" has been introduced to account for rapid improvements beyond the involvement of therapy common factors. As defined by Tang and DeRubeis (1990), critical sessions are those early in the course of therapy that confer a sudden improvement in symptoms or "sudden gains", preceded by "cognitive change" (Tang et al., 2005). Despite initially being a focus of CBT for depression, critical sessions have also been noted to occur in at least seven types of intervention for a range of anxiety disorders and during group therapy (Aderka et al., 2011; Norton et al., 2010). Third, it is possible that there is a subset of patients that may hold certain characteristics that predispose them to rapid improvements. This includes but is not limited to high levels of motivation and readiness to change (Boswell et al., 2012; Lambert & Anderson, 1996). Likewise, there may be therapist factors that can contribute to rapid responses, many of which define a "Super Shrink" as described by Miller and colleagues (2018). These effective therapists are said to have high levels of expertise, be sensitive to feedback and engage with continued development, resulting in a high proportion of Rapid Responders on their caseloads (Hansen et al., 2015). However, meta-analyses have found that therapist factors such as these only account for 5% of variability in outcomes and therefore, cannot solely account for the Rapid Responder class (Baldwin & Imel, 2013). Finally, it is possible that Rapid Responders are a consequence of a general regression to the mean, whereby initial extreme baseline scores (those that are very high or very low) on symptom measures may in fact be a momentary deviation from scores that would otherwise have been closer to the mean score, meaning improvements may not reflect the effect of therapy, but may instead have been due to chance (Nordberg et al. 2014).

The review also identified a smaller subgroup of participants who improved over the course of therapy, albeit during later sessions and were defined as Delayed Responders. This class was the least commonly reported Responder trajectory with around half as many examples as the Rapid Responders counterpart. Research and theory attempting to explain this pattern of symptom change is sparce, perhaps because it not a common feature of GCM study findings. Given the possible reasons for rapid responses, it stands to reason that those with delayed improvements may share some of these, such as the positive impact of non-specific factors or cognitive change, yet simply later in the course of therapy. Furthermore, late improvements may reflect a more complex presentation of a CMD that requires a greater number of sessions to address its symptoms, resulting in a later emergence of improvements (Saunders et al., 2019). In particular, this may apply when there is significant co-morbidity with one presenting difficult serving as a barrier to improvements in another. In such circumstances, early sessions may require a focus on reducing low mood symptoms before the anxiety would be expected to improve, or vice versa, and with this, a delayed response trajectory may be observed. Finally, delayed responses may reflect significant barriers to behavioural change via the medium of psychological intervention, which may take time to resolve. These may include limitations in one's capability, motivation or opportunity to change, requiring transient attention of the therapist and problem-solving before any symptom improvements are noted (Michie et al., 2011). Similarly, delayed responses may also signify patients who require a 'preparatory phase' of cognitive techniques or trauma 'stabilisation', prior to the behavioural components of therapy, such as exposure, which generally confer the most benefits (Bicanic et al., 2015; Gillihan et al., 2013).

Unfortunately, not all patients respond to psychological intervention and some get worse. These individuals have been characterised by the Non-responder and Deteriorator classes identified by a multitude of studies included in the review, with membership usually confined to a minority of participants. It is important to note that there was considerable variation in how Non-Responders were described throughout the literature, with terms such as ‘chronic’ or ‘residual’ symptoms, ‘partial’ or ‘limited’ response and ‘treatment-resistant’ used to describe the pattern of no change that was ultimately shared. According to Gloster and colleagues (2020), these terms suggest a differing viewpoint of the authors as to whether the intervention or the patient was responsible for the lack of improvement. There is also a limited consensus regarding the definition of a ‘response’ as opposed to a ‘non-response’, with some authors choosing to frame very small changes on clinical measures as an improvement, which would otherwise be experienced as a non-response by the patient or perceived as one by the therapist. Acknowledging this, a small number of Responder classes that resembled a partial or limited response were re-classified as Non-responders, using the greater part of Non-responder trajectories as a benchmark, as well as indicators of reliable change (Jacobson & Traux, 1991). The reasons for being a Non-responder are varied, with many yet to be realised. However, they will most likely relate to an interaction of treatment, patient, therapist, service and wider political, social and economic factors as depicted by Figure 3 (Harvery & Gumport, 2016). Alternatively, a patient may subjectively feel better, but outcome measures have failed to capture this (Kounali et al., 2020).

Figure 3

Factors Hypothesised to Influence Trajectory Class Membership



Considering the intervention offered, non-response may occur when there is a mismatch between the therapy offered and the targeted CMD symptoms, or when misdiagnosis has transpired (Taylor et al., 2012). This is of particular relevance for protocolised CBT interventions, which are designed and evaluated for CMDs under the premise of a problem-specific formulation. Alternatively, the right intervention may be provided but delivered with poor adherence to its theoretical underpinnings or the therapist might lack the qualities required for effective treatment delivery, as previously outlined by Miller and colleagues (2018). Meanwhile, the patient may not engage with between session tasks or self-help techniques, be subject to socio-economic issues (e.g., hostile living environments) that maintain distress or have a CMD presentation that is highly complex or severe (Taylor et al., 2012). More recently, attention has been paid to service factors, whereby long wait lists and greater deprivation in the catchment areas of the service have been associated with poorer outcomes for the patient (Clark et al., 2018). The class of Deteriorators are likely to share many of these factors as a reason for an increase in symptoms during a course of therapy, viewed as having failed to impede the natural progression of CMD symptoms. However, this instance of deterioration is distinct from an increase in symptoms because of an intervention considered to be unhelpful or even, harmful (Dimidjian & Hollon, 2010). To date, distinguishing these forms of deterioration has proved challenging, yet the latter is implied when the person is free of life events that might otherwise exacerbate symptoms and when measurement error has been ruled out (Wampold & Imel, 2015). Research into the mechanisms of natural and harmful treatment deterioration is scarce, although some case studies have found that some patients attribute deterioration to the anticipation of the termination of therapy (Bloch-Elkouby et al., 2019).

In review of the 45 studies grouped by the CMD symptom of interest, the proportion of trajectory classes identified and the percentage of the sample that belonged to each class were consistent across participants with either depression, PTSD or a range of anxiety disorders. This supports the view that psychological intervention is effective for all CMDs, with a minority experiencing less favourable outcomes (Cuipers, 2019). In addition, studies that explored the trajectories of two different symptoms experienced by the same sample using separate GCM models generally reported the same trajectories (Batterham et al., 2019; Crane et al., 2019; Galovski et al., 2016; Saunders et al., 2019; Sunderland et al., 2020). Moreover, some studies found that there was an overlap in class membership of different symptoms. In other words, if a patient was a Non-responder on the basis of anxiety symptoms, they were also likely to be a Non-responder in terms of depression symptoms. One study comprised of veterans with PTSD brought into question the importance of etiology, underlying mechanisms of change and their association with trajectory classes. As part of a dismantling trial, this study found that when the trauma processing aspect of Cognitive Processing Therapy (CPT) was removed, leaving solely cognitive components, the membership of the non-responder class increased (Stein et al.., 2012). Consequently, it seems that for some CMDs at least, the mechanism of change of an intervention is an important contributor of symptom improvement.

The number of patient-level variables found to be associated with class membership was substantial. Therefore, those that were more frequently associated or uniformly reported are discussed as follows. First, it was apparent that those with greater baseline severity of CMD symptoms were less likely to benefit from psychological intervention and belong the Non-responder class, as were those with an earlier onset; a finding echoed by a recent study (Buckman et al., 2021). This is broadly consistent with the finding that high levels of functional impairment also predicted membership of this class, since greater symptom severity is often accompanied by significant impairment (Amati et al., 2018). Moreover, participants with greater risk histories, defined by multiple suicide attempts, and comorbidity, in terms of additional diagnoses or symptoms of another CMD, were more likely to be Non-responders. This finding may also extend to physical health comorbidities, with a greater number of conditions, or specific diagnoses such as fibromyalgia, found to be associated with non-response. Aside from the widely recognised link between poorer physical health and unfavorable mental health outcomes, research has highlighted that many therapists lack confidence when working with these populations and find it challenging to adapt interventions or formulations to accommodate for the condition. Meanwhile, many patients find that their engagement or motivation is hindered by poor physical health or find that their physical and mental healthcare are not integrated in a way that would be conducive to CMD symptom improvements (Carroll et al., 2021).

It was commonly reported that those with less perceived or objective social support and those that viewed themselves as a burden tended to be Non-responders. Informal support derived from family, friends, close others or communities has been shown to not only aid formal interventions but also contribute to symptom improvements by its own accord for some individuals (Brown et al., 2014). This was reiterated by a recent meta-analysis, which found that those with severely limited social support had worse depression outcomes than those with adequate social support (Buckman et al., 2021). Furthermore, interventions such as IPT, which emphasise inter-personal processes in the development and maintenance of CMDs, draw on social support as a mechanism of change to alleviate symptoms (Lipsitz & Markowtiz, 2013). Therefore, it is logical that better social support contributes to a more favourable response to intervention.

The role of personality in predicting therapy responses was also explored, with higher levels neuroticism and introversion highlighted as predictors of the Non-responder class. In explanation of this, it is possible that unlike extroverts, introverts do not benefit from an enthusiasm for therapy and a willingness to express themselves during it (Forsell & Åström, 2012). Meanwhile, a high degree of neuroticism may make someone more prone to negative affect and reduce their capacity to tolerate distress, with a reliance on emotion-focused coping; another predictor of the non-response class (Ogrodniczuk et al., 2005). Whilst many studies considered demographic factors as possible predictors for class membership, the results were mostly inconsistent, although Responders tended to be younger and in employment as opposed to

Non-responders. Decline in cognitive flexibility due to ageing has been noted as a possible reason for a lesser response to therapy among older adults, particularly since interventions such as CBT depend on it for skill acquisition and application (Johnco et al., 2014).

Finally, individuals who had previous treatment and either a psychological or pharmacological concurrent intervention were more likely to be Non-responders. It is plausible that the need for additional treatment in the past or present reflects a more complex CMD with symptoms less likely to improve. A concurrent psychological intervention may also impede progression due to a conflicting theoretical underpinning. Moreover, those who reported a poorer therapeutic alliance early in the course of sessions or group cohesion generally did not improve, echoing earlier statements regarding the importance of non-specific factors in determining treatment response (Ilardi & Craighead, 1999). Interestingly, people in receipt of high-intensity or CBT interventions were more likely to be Rapid Responders over Delayed Responders. This may reinforce the premise that CBT sessions encourage cognitive change early on, leading to rapid improvements and that the greater competence of high-intensity, compared to low-intensity, therapists is advantageous (Tang et al., 2005; Gyani et al., 2013).

Dewar and colleagues (2020) sought to summarise PTSD symptom trajectories during interventions and their predictors. In line with the current review, many of the participants in the included studies were veterans, although unlike the current review, inpatient studies were not excluded. They found that three trajectory classes were commonly reported across 11 studies: “Responders”, “Non-Responders” and “Sub-clinical Participants”. Responders were found to have greater social support, less combat exposure, less severe PTSD symptom clusters, lower symptoms of depression at baseline and were younger than Non-responders (Dewar et al., 2020). In comparison, the current review identified 12 studies using a GCM approach to identifying

PTSD symptom trajectories during outpatient interventions, whilst Responders and Non-responders remained the most common classes among them. However, Dewar and colleagues (2020) did not report the subclasses of Delayed and Rapid Responders, nor did they highlight Deteriorators. Furthermore, they defined a novel class of subclinical participants by re-classifying those with baseline symptoms that did not meet the clinical threshold of relevant PTSD measures. The predictors of the PTSD Non-responder class were consistent with these previous findings, whereby for PTSD in particular, greater combat exposure and more severe symptom clusters (guilt, avoidance, re-experiencing and hyperarousal) were associated with trajectory class membership.

Limitations

The findings of the review are subject to a number of limitations. Firstly, the samples were predominantly female and white, limiting the generalisability of the findings. Moreover, a concerning number of studies failed to report the ethnic background of their participants at all. To ensure the effectiveness of interventions for people of minority ethnic backgrounds, it is imperative that provisions of care and therapists are sensitive to cultural factors relevant to the development and presentation of CMDs, as well as those that influence the therapeutic alliance (Meyer & Zane, 2013). Because of this, it is feasible that ethnicity may predict trajectory class membership and should therefore, not be overlooked as it has by many studies of this review. Considering the use of the GRoLTS checklist, it is apparent that the quality of GCM research is highly varied, with the majority of studies not fulfilling a number of checklist items, perhaps due to the checklist's fairly recent publication. Therefore, the conclusions of the review should be read with caution due to the inclusion of some studies with sub-optimal reporting standards. The study is also limited given that the number of measurement points utilised for GCM across the studies was highly varied and some only focused on specific phases of treatment, whilst others included the whole course as well as follow-up. Because of this, the ability to make valid comparisons is reduced. As previously noted, there was a correlation between the number of classes reported and study sample size, where more classes were identified as sample size increased; a finding supported by Monte Carlo evaluations of GCM analyses (Kim, 2014; Shader & Beauchaine, 2021). Since most samples of the studies included in the review were relatively small, it is possible that the number of classes were underestimated. The data synthesised by the review was not amenable to meta-analytic methods and therefore, the findings are limited by a narrative approach to synthesis. Narrative reviews are subject to a range of criticisms such as a lack of rigorously described methods and ambiguity in the links between the data and the interpretations that are made. Nonetheless attempts were made to address these, through drawing on aspects of the synthesis without meta-analysis guidance (SWiM; Campbell et al., 2020). Finally, few studies, aside those that were CBT orientated, explored symptom trajectory classes during a sole provision of therapy. Instead, numerous studies offered different types of interventions and did not apply separate GCM models for subsamples of individuals receiving distinct therapies. Because of this, direct comparisons were not possible. Future studies would benefit from utilising such an approach.

Clinical and Research Implications

The findings of the current review have potential clinical and research implications that may improve patient care. Findings suggest that regardless of the type of intervention offered, the symptom class trajectories are mostly comparable. Coupled with broadly equivalent efficacies, this suggests that there is likely to be greater symptom change variability within, rather than between, interventions. In light of this, non-therapy specific factors relevant to the patient, therapist and service delivery are of great importance (Saxon et al., 2017). Therefore, clinicians are advised to move beyond the traditional stance of perceiving certain therapy types as superior, in favour of addressing the crucial question of: “What Works for Whom?” (Roth & Fonagy, 2005). In other words, tailored interventions, which take into account non-therapy specific factors and patient choice, are vital. Thus, outcome feedback systems used during routine outcome monitoring that identify expected intervention response trajectories based on predictors, such as those identified the current review, might be of further benefit for patients, therapists and services alike (Delgadillo et al., 2018). Building on this, it is feasible that future studies, which strive to overcome the previously mentioned limitations of the studies included in this review, could refine these feedback systems to predict differing responses to varying therapies and in doing so, identify those that are most conducive to CMD improvements for the individual referred.

References

Aderka, I. M., Nickerson, A., Bøe, H. S., Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: a meta-analysis. *Journal of Consulting and Clinical Psychology, 80*(1), 93-101. https://doi.org/10.1037/a0026455

Allan, N. P., Gros, D. F., Myers, U. S., Korte, K. J., & Acierno, R. (2017). Predictors and Outcomes of Growth Mixture Modeled Trajectories Across an Exposure-Based PTSD Intervention with Veterans. *Journal of clinical psychology, 73*(9), 1048–1063. https://doi.org/10.1002/jclp.22408

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

Baldwin, S. & Imel, Z. E. (2013). Therapist Effects: Findings and Methods. In Lambert, M. J. (Ed.). *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed., pp. 258-297). Wiley.

Bandelow, B., Reitt, M., Röver, C., Michaelis, S., Görlich, Y., & Wedekind, D. (2015). Efficacy of treatments for anxiety disorders: a meta-analysis. *International clinical psychopharmacology, 30*(4), 183–192. https://doi.org/10.1097/YIC.0000000000000078

Barkham, M., Saxon, D., Hardy, G. E., Bradburn, M., Galloway, D., Wickramasekera, N., Keetharuth, A. D., Bower, P., King, M., Elliott, R., Gabriel, L., Kellett, S., Shaw, S., Wilkinson, T., Connell, J., Harrison, P., Ardern, K., Bishop-Edwards, L., Ashley, K., Ohlsen, S., … Brazier, J. E. (2021). Person-centred experiential therapy versus cognitive behavioural therapy delivered in the English Improving Access to Psychological Therapies service for the treatment of moderate or severe depression (PRaCTICED): a pragmatic, randomised, non-inferiority trial. *The Lancet Psychiatry, 8*(6), 487–499. https://doi.org/10.1016/S2215-0366(21)00083-3

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: toward practice-based evidence in the psychological therapies. *Clinical Outcomes in Routine Evaluation-Outcome Measures. Journal of consulting and clinical psychology, 69*(2), 184–196. https://doi.org/10.1037//0022-006x.69.2.184

Batterham, P. J., Christensen, H., Mackinnon, A. J., Gosling, J. A., Thorndike, F. P., Ritterband, L. M., Glozier, N., & Griffiths, K. M. (2017). Trajectories of change and long-term outcomes in a randomised controlled trial of internet-based insomnia treatment to prevent depression. *BJPsych open, 3*(5), 228–235. https://doi.org/10.1192/bjpo.bp.117.005231

Batterham, P. J., van Spijker, B., Mackinnon, A. J., Calear, A. L., Wong, Q., & Christensen, H. (2019). Consistency of trajectories of suicidal ideation and depression symptoms: Evidence from a randomized controlled trial. *Depression and anxiety, 36*(4), 321–329. https://doi.org/10.1002/da.22863

Bei, B., Asarnow, L. D., Krystal, A., Edinger, J. D., Buysse, D. J., & Manber, R. (2018). Treating insomnia in depression: Insomnia related factors predict long-term depression trajectories. *Journal of consulting and clinical psychology, 86*(3), 282–293. https://doi.org/10.1037/ccp0000282

Bicanic, I., de Jongh, A., & Ten Broeke, E. (2015). Stabilisation in trauma treatment: necessity or myth?. *Tijdschrift voor psychiatrie, 57*(5), 332–339.

Bloch-Elkouby, S., Eubanks, C. F., Knopf, L., Gorman, B. S., & Muran, J. C. (2019). The Difficult Task of Assessing and Interpreting Treatment Deterioration: An Evidence-Based Case Study. *Frontiers in psychology, 10*, 1180. https://doi.org/10.3389/fpsyg.2019.01180

Boswell, J. F., Sauer-Zavala, S. E., Gallagher, M. W., Delgado, N. K., & Barlow, D. H. (2012). Readiness to change as a moderator of outcome in transdiagnostic treatment. *Journal of the Society for Psychotherapy Research, 22*(5), 570–578. https://doi.org/10.1080/10503307.2012.688884

Bray, R. M., Engel, C. C., Williams, J., Jaycox, L. H., Lane, M. E., Morgan, J. K., & Unützer, J. (2016). Posttraumatic Stress Disorder in U.S. Military Primary Care: Trajectories and Predictors of One-Year Prognosis. *Journal of traumatic stress, 29*(4), 340–348. https://doi.org/10.1002/jts.22119

Brown, J., Evans-Lacko, S., Aschan, L., Henderson, M. J., Hatch, S. L., & Hotopf, M. (2014). Seeking informal and formal help for mental health problems in the community: a secondary analysis from a psychiatric morbidity survey in South London. *BMC psychiatry, 14*(1), 275. https://doi.org/10.1186/s12888-014-0275-y

Buckman, J., Saunders, R., O'Driscoll, C., Cohen, Z. D., Stott, J., Ambler, G., Gilbody, S., Hollon, S. D., Kendrick, T., Watkins, E., Wiles, N., Kessler, D., Chari, N., White, I. R., Lewis, G., & Pilling, S. (2021). Is social support pre-treatment associated with prognosis for adults with depression in primary care?. *Acta psychiatrica Scandinavica, 143*(5), 392–405. https://doi.org/10.1111/acps.13285

Buckman, J. E. J., Cohen, Z. D., O'Driscoll, C., Fried, E. I., Saunders, R., Ambler, G., DeRubeis, R. J., Gilbody, S., Hollon, S. D., Kendrick, T., Watkins, E., Eley, T. C., Peel, A. J., Ranyer, C., Kessler, D., Wiles, N., Lewis, G., Pilling, S. (2021). Predicting prognosis for adults with depression using individual symptom data: a comparison of modelling approaches. *Psychological Medicine*, 1-11. https://doi.org/10.1017/S0033291721001616

Burlingame, G. M., Seebeck, J. D., Janis, R. A., Whitcomb, K. E., Barkowski, S., Rosendahl, J., & Strauss, B. (2016). Outcome differences between individual and group formats when identical and nonidentical treatments, patients, and doses are compared: A 25-year meta-analytic perspective. *Psychotherapy, 53*(4), 446–461. https://doi.org/10.1037/pst0000090

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ (Clinical research ed.), 368*. https://doi.org/10.1136/bmj.l6890

Carroll, S., Moss-Morris, R., Hulme, K., & Hudson, J. (2021). Therapists' perceptions of barriers and facilitators to uptake and engagement with therapy in long-term conditions. *British journal of health psychology, 26*(2), 307–324. https://doi.org/10.1111/bjhp.12475

Clark D. M. (2018). Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program. *Annual review of clinical psychology, 14*, 159–183. https://doi.org/10.1146/annurev-clinpsy-050817-084833

Clark, D. M., Canvin, L., Green, J., Layard, R., Pilling, S., & Janecka, M. (2018). Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. *Lancet, 391*(10121), 679–686. https://doi.org/10.1016/S0140-6736(17)32133-5

Cohen, Z. D., & DeRubeis, R. J. (2018). Treatment Selection in Depression. *Annual review of clinical psychology, 14*, 209–236. https://doi.org/10.1146/annurev-clinpsy-050817-084746

Crane, T. E., Badger, T. A., Sikorskii, A., Segrin, C., Hsu, C. H., & Rosenfeld, A. G. (2019). Trajectories of Depression and Anxiety in Latina Breast Cancer Survivors. *Oncology nursing forum, 46*(2), 217–227. https://doi.org/10.1188/19.ONF.217-227

Cuijpers, P. (2017). Four decades of outcome research on psychotherapies for adult depression: An overview of a series of meta-analyses. *Canadian Psychology, 58*(1), 7–19. https://doi.org/10.1037/cap0000096

Cuijpers, P. (2019). Targets and outcomes of psychotherapies for mental disorders: an overview. *World Psychiatry, 18*, 276-285. https://doi.org/10.1002/wps.20661

Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian journal of psychiatry, 58*(7), 376–385. https://doi.org/10.1177/070674371305800702

Cuijpers, P., van Lier, P. A., van Straten, A., & Donker, M. (2005). Examining differential effects of psychological treatment of depressive disorder: an application of trajectory analyses. *Journal of affective disorders, 89*(1-3), 137–146. https://doi.org/10.1016/j.jad.2005.09.001

Curran, P. J., Obeidat, K., & losardo, D. (2010). Twelve Frequently Asked Questions About Growth Curve Modeling. *Journal of Cognitive Development, 11*(2), 121-136. https://doi.org/10.1080/15248371003699969

Delgadillo, J., de Jong, K., Lucock, M., Lutz, W., Rubel, J., Gilbody, S., Ali, S., Aguirre, E., Appleton, M., Nevin, J., O'Hayon, H., Patel, U., Sainty, A., Spencer, P., & McMillan, D. (2018). Feedback-informed treatment versus usual psychological treatment for depression and anxiety: a multisite, open-label, cluster randomised controlled trial. *The lancet. Psychiatry, 5*(7), 564–572. https://doi.org/10.1016/S2215-0366(18)30162-7

DeRubeis, R. J., Evans, M. D., Hollon, S. D., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1990). How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of consulting and clinical psychology, 58*(6), 862–869. https://doi.org/10.1037//0022-006x.58.6.862

Dewar, M., Paradis, A., & Fortin, C. A. (2020). Identifying Trajectories and Predictors of Response to Psychotherapy for Post-Traumatic Stress Disorder in Adults: A Systematic Review of Literature. *The Canadian Journal of Psychiatry, 65*(2), 71–86. https://doi.org/10.1177/0706743719875602

Dimidjian, S., & Hollon, S. D. (2010). How would we know if psychotherapy were harmful? *American Psychologist, 65*(1), 21–33. https://doi.org/10.1037/a0017299

Eftekhari, A., Crowley, J. J., Mackintosh, M. A., & Rosen, C. S. (2020). Predicting treatment dropout among veterans receiving prolonged exposure therapy. *Psychological trauma: theory, research, practice and policy, 12*(4), 405–412. https://doi.org/10.1037/tra0000484

Fanous, M. & Daniels, J. (2020). What are the recovery and attrition outcomes for group CBT and individual CBT for Generalised Anxiety Disorder in an IAPT Service? An exploratory study. *The Cognitive Behaviour Therapist, 13*(15). https://doi.org/10.1017/S1754470X20000045

Fletcher, S., Elklit, A., Shevlin, M., & Armour, C. (2021). Predictors of PTSD Treatment Response Trajectories in a Sample of Childhood Sexual Abuse Survivors: The Roles of Social Support, Coping, and PTSD Symptom Clusters. *Journal of interpersonal violence, 36*(3-4), 1283–1307. https://doi.org/10.1177/0886260517741212

Flood, N., Page, A., & Hooke, G. (2019). A comparison between the clinical significance and growth mixture modelling early change methods at predicting negative outcomes. *Psychotherapy research: journal of the Society for Psychotherapy Research, 29*(7), 947–958. https://doi.org/10.1080/10503307.2018.1469803

Forsell, L. M., & Åström, J. A. (2012). An Analysis of Resistance to Change Exposed in Individuals' Thoughts and Behaviors. *Comprehensive Psychology*. https://doi.org/10.2466/09.02.10.CP.1.17

Fournier, J. C., DeRubeis, R. J., Shelton, R. C., Hollon, S. D., Amsterdam, J. D., & Gallop, R. (2009). Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *Journal of Consulting and Clinical Psychology, 77*(4), 775–787. https://doi.org/10.1037/a0015401

Frankfurt, S., Frazier, P., Litz, B. T., Schnurr, P. P., Orazem, R. J., Gravely, A., & Sayer, N. (2019). Online expressive writing intervention for reintegration difficulties among veterans: Who is most likely to benefit?. *Psychological trauma: theory, research, practice and policy, 11*(8), 861–868. https://doi.org/10.1037/tra0000462

Galovski, T. E., Harik, J. M., Blain, L. M., Farmer, C., Turner, D., & Houle, T. (2016). Identifying patterns and predictors of PTSD and depressive symptom change during cognitive processing therapy. *Cognitive Therapy and Research, 40*(5), 617–626. https://doi.org/10.1007/s10608-016-9770-4

Gillihan, S. J., Williams, M. T., Malcoun, E., Yadin, E., & Foa, E. B. (2012). Common Pitfalls in Exposure and Response Prevention (EX/RP) for OCD. *Journal of obsessive-compulsive and related disorders, 1*(4), 251–257. https://doi.org/10.1016/j.jocrd.2012.05.002

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

Guidi, J., Brakemeier, E. L., Bockting, C., Cosci, F., Cuijpers, P., Jarrett, R. B., Linden, M., Marks, I., Peretti, C. S., Rafanelli, C., Rief, W., Schneider, S., Schnyder, U., Sensky, T., Tomba, E., Vazquez, C., Vieta, E., Zipfel, S., Wright, J. H., & Fava, G. A. (2018). Methodological Recommendations for Trials of Psychological Interventions. *Psychotherapy and psychosomatics, 87*(5), 276–284. https://doi.org/10.1159/000490574

Gyani, A., Shafran, R., Layard, R., & Clark, D. M. (2013). Enhancing recovery rates: lessons from year one of IAPT. *Behaviour research and therapy, 51*(9), 597–606. https://doi.org/10.1016/j.brat.2013.06.004

Haas, E., Hill, R. D., Lambert, M. J., & Morrell, B. (2002). Do early responders to psychotherapy maintain treatment gains. *Journal of Clinical Psychology, 58*(9), 1157–1172. https://doi.org/10.1002/jclp.10044

Hansen, B. P., Lambert, M. J., & Vlass, E. N. (2015). Sudden gains and sudden losses in the clients of a "supershrink": 10 case studies. *Pragmatic Case Studies in Psychotherapy, 11*(3), 154–201. https://doi.org/10.14713/pcsp.v11i3.1915

Harvey, A. G., & Gumport, N. B. (2015). Evidence-based psychological treatments for mental disorders: modifiable barriers to access and possible solutions. *Behaviour research and therapy, 68*, 1–12. https://doi.org/10.1016/j.brat.2015.02.004

Heckman, T. G., Heckman, B. D., Anderson, T., Bianco, J. A., Sutton, M., & Lovejoy, T. I. (2017). Common Factors and Depressive Symptom Relief Trajectories in Group Teletherapy for Persons Ageing with HIV. *Clinical psychology & psychotherapy, 24*(1), 139–148. https://doi.org/10.1002/cpp.1989

Hull, T. D., Malgaroli, M., Connolly, P. S., Feuerstein, S., & Simon, N. M. (2020). Two-way messaging therapy for depression and anxiety: longitudinal response trajectories. *BMC psychiatry, 20*(1), 297. https://doi.org/10.1186/s12888-020-02721-x

Ilardi, S. S., & Craighead, W. E. (1999). Rapid early response, cognitive modification, and nonspecific factors in cognitive behavior therapy for depression: A reply to Tang and DeRubeis. *Clinical Psychology: Science and Practice, 6*(3), 295–299. https://doi.org/10.1093/clipsy.6.3.295

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of consulting and clinical psychology, 59*(1), 12–19. https://doi.org/10.1037//0022-006x.59.1.12

Joesch, J. M., Golinelli, D., Sherbourne, C. D., Sullivan, G., Stein, M. B., Craske, M. G., & Roy-Byrne, P. P. (2013). Trajectories of change in anxiety severity and impairment during and after treatment with evidence-based treatment for multiple anxiety disorders in primary care. *Depression and anxiety, 30*(11), 1099–1106. https://doi.org/10.1002/da.22149

Johnco, C., Wuthrich, V. M., & Rapee, R. M. (2014). The influence of cognitive flexibility on treatment outcome and cognitive restructuring skill acquisition during cognitive behavioural treatment for anxiety and depression in older adults: Results of a pilot study. *Behaviour research and therapy, 57*, 55–64. https://doi.org/10.1016/j.brat.2014.04.005

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1-27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432

Kendrick, T., & Pilling, S. (2012). Common mental health disorders--identification and pathways to care: NICE clinical guideline. *The British journal of general practice: the journal of the Royal College of General Practitioners, 62*(594), 47–49. https://doi.org/10.3399/bjgp12X616481

Kim S. Y. (2014). Determining the Number of Latent Classes in Single- and Multi-Phase Growth Mixture Models. *Structural equation modeling: a multidisciplinary journal, 21*(2), 263–279. https://doi.org/10.1080/10705511.2014.882690

Koffmann A. (2018). Has growth mixture modeling improved our understanding of how early change predicts psychotherapy outcome?. *Psychotherapy research: journal of the Society for Psychotherapy Research, 28*(6), 829–841. https://doi.org/10.1080/10503307.2017.1294771

Kounali, D., Button, K. S., Lewis, G., Gilbody, S., Kessler, D., Araya, R., Duffy, L., Lanham, P., Peters, T. J., Wiles, N., & Lewis, G. (2020). How much change is enough? Evidence from a longitudinal study on depression in UK primary care. *Psychological medicine,* 1–8. https://doi.org/10.1017/S0033291720003700

Lambert, M. J., & Anderson, E. M. (1996). Assessment for the time-limited psychotherapies. *American Psychiatric Press Review of Psychiatry, 15*, 23–42.

Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D. A., Nielsen, S. L., & Hawkins, E. J. (2001). The effects of providing therapists with feedback on patient progress during psychotherapy: are outcomes enhanced?. *Psychotherapy research: journal of the Society for Psychotherapy Research, 11*(1), 49–68. https://doi.org/10.1080/713663852

Lewis, C. C., Simons, A. D., & Kim, H. K. (2012). The role of early symptom trajectories and pretreatment variables in predicting treatment response to cognitive behavioral therapy. *Journal of consulting and clinical psychology, 80*(4), 525–534. https://doi.org/10.1037/a0029131

Lipsitz, J. D., & Markowitz, J. C. (2013). Mechanisms of change in interpersonal therapy (IPT). *Clinical psychology review, 33*(8), 1134–1147. https://doi.org/10.1016/j.cpr.2013.09.002

Litz, B. T., Berke, D. S., Kline, N. K., Grimm, K., Rusowicz-Orazem, L., Resick, P. A., Foa, E. B., Wachen, J. S., McLean, C. P., Dondanville, K. A., Borah, A. M., Roache, J. D., Young-McCaughan, S., Yarvis, J. S., Mintz, J., & Peterson, A. L. (2019). Patterns and predictors of change in trauma-focused treatments for war-related posttraumatic stress disorder. *Journal of consulting and clinical psychology, 87*(11), 1019–1029. https://doi.org/10.1037/ccp0000426

Lukaschek, K., Hiller, T. S, Schumacher, U., Teismann, T., Breitbart, J., Brettschneider, C., König, H., Margraf, J., & Gensichen, J. (2019). Symptom trajectories in patients with panic disorder in a primary care intervention: Results from a randomized controlled trial (PARADISE). *Scientific Reports, 9*(7170). https://doi.org/10.1038/s41598-019-43487-x

Lutz, W., Arndt, A., Rubel, J., Berger, T., Schröder, J., Späth, C., Meyer, B., Greiner, W., Gräfe, V., Hautzinger, M., Fuhr, K., Rose, M., Nolte, S., Löwe, B., Hohagen, F., Klein, J. P., & Moritz, S. (2017). Defining and Predicting Patterns of Early Response in a Web-Based Intervention for Depression. *Journal of medical Internet research, 19*(6). https://doi.org/10.2196/jmir.7367

Lutz, W., Hofmann, S. G., Rubel, J., Boswell, J. F., Shear, M. K., Gorman, J. M., Woods, S. W., & Barlow, D. H. (2014). Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *Journal of consulting and clinical psychology, 82*(2), 287–297. https://doi.org/10.1037/a0035535

Lutz, W., Prinz, J. N., Schwartz, B., Paulick, J., Schoenherr, D., Deisenhofer, A. K., Terhürne, P., Boyle, K., Altmann, U., Strauß, B., Rafaeli, E., Atzil-Slonim, D., Bar-Kalifa, E., & Rubel, J. (2020). Patterns of early change in interpersonal problems and their relationship to nonverbal synchrony and multidimensional outcome. *Journal of counseling psychology, 67*(4), 449–461. https://doi.org/10.1037/cou0000376

Lutz, W., Stulz, N., & Köck, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *Journal of affective disorders, 118*(1-3), 60–68. https://doi.org/10.1016/j.jad.2009.01.019

Malgaroli, M., Hull, T. D., Wiltsey Stirman, S., & Resick, P. (2020). Message Delivery for the Treatment of Posttraumatic Stress Disorder: Longitudinal Observational Study of Symptom Trajectories. *Journal of medical Internet research, 22*(4). https://doi.org/10.2196/15587

Marcelle, E. T., Nolting, L., Hinshaw, S. P., & Aguilera, A. (2019). Effectiveness of a Multimodal Digital Psychotherapy Platform for Adult Depression: A Naturalistic Feasibility Study. *JMIR mHealth and uHealth, 7*(1). https://doi.org/10.2196/10948

McDevitt-Petrovic, O., Shevlin, M., & Kirby, K. (2020). Modelling changes in anxiety and depression during low-intensity cognitive behavioural therapy: An application of growth mixture models. *The British journal of clinical psychology, 59*(2), 169–185. https://doi.org/10.1111/bjc.12237

McManus S, Bebbington P, Jenkins R, Brugha T. (eds.) (2016) *Mental health and wellbeing in England: Adult Psychiatric Morbidity Survey 2014*. NHS Digital. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/556596/apms-2014-full-rpt.pdf

Meyer, O. L., & Zane, N. (2013). The Influence of Race and Ethnicity in Clients' Experiences of Mental Health Treatment. *Journal of community psychology, 41*(7), 884–901. https://doi.org/10.1002/jcop.21580

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation science, 6*(42). https://doi.org/10.1186/1748-5908-6-42

Miller, S. D., Hubble, M. A., & Chow, D. (2018). The question of expertise in psychotherapy. *Journal of Expertise, 1*(2). https://www.journalofexpertise.org/articles/JoE_2018_1_2_MillerHubbardChow_earlyview.pdf

Moggia, D., Lutz, W., Arndt, A., & Feixas, G. (2020). Patterns of change and their relationship to outcome and follow-up in group and individual psychotherapy for depression. *Journal of consulting and clinical psychology, 88*(8), 757–773. https://doi.org/10.1037/ccp0000562

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7). https://doi.org/10.1371/journal.pmed.1000097

Murphy, D., & Smith, K. V. (2018). Treatment Efficacy for Veterans with Posttraumatic Stress Disorder: Latent Class Trajectories of Treatment Response and Their Predictors. *Journal of traumatic stress, 31*(5), 753–763. https://doi.org/10.1002/jts.22333

NHS Digital (2019). *Psychological Therapies, Annual report on the use of IAPT services 2018-19*. https://digital.nhs.uk/data-and-information/publications/statistical/psychological-therapies-annual-reports-on-the-use-of-iapt-services/annual-report-2018-19

NICE (2009). *Depression in adults: recognition and management Clinical guideline [CG90].* https://www.nice.org.uk/guidance/cg90/chapter/Recommendations

NICE (2014). *Anxiety disorders Quality standard [QS53].* https://www.nice.org.uk/guidance/qs53

NICE (2018). *Post-traumatic stress disorder NICE guideline [NG116].* https://www.nice.org.uk/guidance/ng116

Nordberg, S. S., Castonguay, L. G., Fisher, A. J., Boswell, J. F., & Kraus, D. (2014). Validating the rapid responder construct within a practice research network. *Journal of clinical psychology, 70*(9), 886–903. https://doi.org/10.1002/jclp.22077

Norton, P. J., Klenck, S. C., & Barrera, T. L. (2010). Sudden gains during cognitive-behavioral group therapy for anxiety disorders. *Journal of anxiety disorders, 24*(8), 887–892. https://doi.org/10.1016/j.janxdis.2010.06.012

Ogrodniczuk, J. S., Piper, W. E., Joyce, A. S. (2005). The negative effect of alexithymia on the outcome of group therapy for complicated grief: what role might the therapist play?. *Comprehensive Psychiatry, 46*(3), 206-213. https://doi.org/10.1016/j.comppsych.2004.08.005

Owen, J., Adelson, J., Budge, S., Wampold, B., Kopta, M., Minami, T., & Miller, S. (2015). Trajectories of Change in Psychotherapy. *Journal of clinical psychology, 71*(9), 817–827. https://doi.org/10.1002/jclp.22191

Petersen, J. J., Hartig, J., Paulitsch, M. A., Pagitz, M., Mergenthal, K., Rauck, S., Reif, A., Gerlach, F. M., & Gensichen, J. (2018). Classes of depression symptom trajectories in patients with major depression receiving a collaborative care intervention. *PLoS ONE, 13*(9). https://doi.org/10.1371/journal.pone.0202245

Phelps, A. J., Steel, Z., Metcalf, O., Alkemade, N., Kerr, K., O'Donnell, M., Nursey, J., Cooper, J., Howard, A., Armstrong, R., & Forbes, D. (2018). Key patterns and predictors of response to treatment for military veterans with post-traumatic stress disorder: a growth mixture modelling approach. *Psychological medicine, 48*(1), 95–103. https://doi.org/10.1017/S0033291717001404

PROSPERO: International prospective register of systematic reviews. (2020). *A systematic review of common mental disorder symptom trajectories and their predictors during psychological interventions*. CRD42020212497. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020212497

Pybis, J., Saxon, D., Hill, A., & Barkham, M. (2017). The comparative effectiveness and efficiency of cognitive behaviour therapy and generic counselling in the treatment of depression: evidence from the 2nd UK National Audit of psychological therapies. *BMC psychiatry, 17*(1), 215. https://doi.org/10.1186/s12888-017-1370-7

Ram, N., & Grimm, K. J. (2009). Growth Mixture Modeling: A Method for Identifying Differences in Longitudinal Change Among Unobserved Groups. *International Journal of Behavioural Development, 33*(6), 565-576. https://doi.org/10.1177/0165025409343765

Ram, N., & Grimm, K. J. (2009). Growth Mixture Modeling: A Method for Identifying Differences in Longitudinal Change Among Unobserved Groups. *International journal of behavioral development, 33*(6), 565–576. https://doi.org/10.1177/0165025409343765

Richards, D., Enrique, A., Eilert, N., Franklin, M., Palacios, J., Duffy, D., Earley, C., Chapman, J., Jell, G., Sollesse, S., & Timulak, L. (2020). A pragmatic randomized waitlist-controlled effectiveness and cost-effectiveness trial of digital interventions for depression and anxiety. *Npj Digital Medicine, 3*(85). https://doi.org/10.1038/s41746-020-0293-8

Roth, A., Fonagy, P. (2005). *What works for whom: A critical review of psychotherapy research* (2nd ed.). Guilford Publications.

Rubel, J., Lutz, W., Kopta, S. M., Köck, K., Minami, T., Zimmermann, D., & Saunders, S. M. (2015). Defining early positive response to psychotherapy: An empirical comparison between clinically significant change criteria and growth mixture modeling. *Psychological assessment, 27*(2), 478–488. https://doi.org/10.1037/pas0000060

Santacana, M., Arias, B., Mitjans, M., Bonillo, A., Montoro, M., Rosado, S., Guillamat, R., Vallès, V., Pérez, V., Forero, C. G., & Fullana, M. A. (2016). Predicting Response Trajectories during Cognitive-Behavioural Therapy for Panic Disorder: No Association with the BDNF Gene or Childhood Maltreatment. *PloS one, 11*(6). https://doi.org/10.1371/journal.pone.0158224

Saunders, R., Buckman, J., Cape, J., Fearon, P., Leibowitz, J., & Pilling, S. (2019). Trajectories of depression and anxiety symptom change during psychological therapy. *Journal of affective disorders, 249*, 327–335. https://doi.org/10.1016/j.jad.2019.02.043

Saunders, R., Buckman, J., & Pilling, S. (2020). Latent variable mixture modelling and individual treatment prediction. *Behaviour research and therapy, 124*, 103505. https://doi.org/10.1016/j.brat.2019.103505

Saxon, D., Barkham, M., Foster, A., & Parry, G. (2017). The Contribution of Therapist Effects to Patient Dropout and Deterioration in the Psychological Therapies. *Clinical psychology & psychotherapy, 24*(3), 575–588. https://doi.org/10.1002/cpp.2028

Schardt, C., Adams, M. B., Owens, T., Keitz, S., & Fontelo, P. (2007). Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC medical informatics and decision making, 7*(16). https://doi.org/10.1186/1472-6947-7-16

Schumm, J. A., Walter, K. H., & Chard, K. M. (2013). Latent class differences explain variability in PTSD symptom changes during cognitive processing therapy for veterans. *Psychological Trauma: Theory, Research, Practice, and Policy, 5*(6), 536–544. https://doi.org/10.1037/a0030359

Shader, T. M., & Beauchaine, T. P. (2021). A Monte Carlo evaluation of growth mixture modeling. *Development and psychopathology*, 1–14. https://doi.org/10.1017/S0954579420002230

Siddique, J., Chung, J. Y., Brown, C. H., & Miranda, J. (2012). Comparative effectiveness of medication versus cognitive-behavioral therapy in a randomized controlled trial of low-income young minority women with depression. *Journal of consulting and clinical psychology, 80*(6), 995–1006. https://doi.org/10.1037/a0030452

Smits, D., Stinckens, N., Luyckx, K., & Claes, L. (2016). Early symptom change in adult outpatients: Relationship with patient characteristics and therapeutic alliance. *Psychology and psychotherapy, 89*(4), 402–417. https://doi.org/10.1111/papt.12086

Steel, Z., Marnane, C., Iranpour, Chey, T., Jackson, J. W., Patel, V., & Silove, D. (2014). The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *International Journal of Epidemiology, 43*(2), 476–493. https://doi.org/10.1093/ije/dyu038

Stefanopoulou, E., Lewis, D., Taylor, M., Broscombe, J., & Larkin, J. (2019). Digitally Delivered Psychological Interventions for Anxiety Disorders: a Comprehensive Review. *The Psychiatric quarterly, 90*(1), 197–215. https://doi.org/10.1007/s11126-018-9620-5

Stein, N. R., Dickstein, B. D., Schuster, J., Litz, B. T., & Resick, P. A. (2012). Trajectories of response to treatment for posttraumatic stress disorder. *Behavior therapy, 43*(4), 790–800. https://doi.org/10.1016/j.beth.2012.04.003

Stulz, N., & Lutz, W. (2007). Multidimensional patterns of change in outpatient psychotherapy: the phase model revisited. *Journal of clinical psychology, 63*(9), 817–833. https://doi.org/10.1002/jclp.20397

Stulz, N., Lutz, W., Leach, C., Lucock, M., & Barkham, M. (2007). Shapes of early change in psychotherapy under routine outpatient conditions. *Journal of Consulting and Clinical Psychology, 75*(6), 864–874. https://doi.org/10.1037/0022-006X.75.6.864

Stulz, N., Thase, M. E., Klein, D. N., Manber, R., & Crits-Christoph, P. (2010). Differential effects of treatments for chronic depression: a latent growth model reanalysis. *Journal of consulting and clinical psychology, 78*(3), 409–419. https://doi.org/10.1037/a0019267

Sunderland, M., Wong, N., Hilvert-Bruce, Z., & Andrews, G. (2012). Investigating trajectories of change in psychological distress amongst patients with depression and generalised anxiety disorder treated with internet cognitive behavioural therapy. *Behaviour research and therapy, 50*(6), 374–380. https://doi.org/10.1016/j.brat.2012.03.005

Tang, T. Z., DeRubeis, R. J., Beberman, R., & Pham, T. (2005). Cognitive changes, critical sessions, and sudden gains in cognitive-behavioral therapy for depression. *Journal of consulting and clinical psychology, 73*(1), 168–172. https://doi.org/10.1037/0022-006X.73.1.168

Taylor, S., Abramowitz, J. S., & McKay, D. (2012). Non-adherence and non-response in the treatment of anxiety disorders. *Journal of anxiety disorders, 26*(5), 583–589. https://doi.org/10.1016/j.janxdis.2012.02.010

Thibodeau, M. A., Quilty, L. C., De Fruyt, F., De Bolle, M., Rouillon, F., & Bagby, R. M. (2015). Latent classes of nonresponders, rapid responders, and gradual responders in depressed outpatients receiving antidepressant medication and psychotherapy. *Depression and anxiety, 32(*3), 213–220. https://doi.org/10.1002/da.22293

Thornicroft, G., Chatterji, S., Evans-Lacko, S., Gruber, M., Sampson, N., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Andrade, L., Borges, G., Bruffaerts, R., Bunting, B., de Almeida, J. M., Florescu, S., de Girolamo, G., Gureje, O., Haro, J. M., He, Y., Hinkov, H., Karam, E., … Kessler, R. C. (2017). Undertreatment of people with major depressive disorder in 21 countries. *The British journal of psychiatry: the journal of mental science, 210*(2), 119–124. https://doi.org/10.1192/bjp.bp.116.188078

Uher, R., Muthén, B., Souery, D., Mors, O., Jaracz, J., Placentino, A., Petrovic, A., Zobel, A., Henigsberg, N., Rietschel, M., Aitchison, K. J., Farmer, A., & McGuffin, P. (2010). Trajectories of change in depression severity during treatment with antidepressants. *Psychological medicine, 40*(8), 1367–1377. https://doi.org/10.1017/S0033291709991528

van der Nest, G., Passos, V. L., Candel, M. J. J. M., & van Breukelen, G. J. P. (2020). An overview of mixture modelling for latent evolutions in longitudinal data: Modelling approaches, fit statistics and software. *Advances in Life Course Research, 43*. https://doi.org/10.1016/j.alcr.2019.100323

van de Schoot, R., Sijbrandij, M., Winter, S. D., Depaoli, S., & Vermunt, J. K. (2017). The GRoLTS-checklist: Guidelines for reporting on latent trajectory studies. *Structural Equation Modeling, 24*(3), 451–467. https://doi.org/10.1080/10705511.2016.1247646

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110*(20), 639–644. https://doi.org/10.2466/02.07.17.PR0.110.2.639-644

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge/Taylor & Francis Group.

Wardenaar, K. J., Conradi, H. J., & de Jonge, P. (2014). Data-driven course trajectories in primary care patients with major depressive disorder. *Depression and anxiety, 31*(9), 778–786. https://doi.org/10.1002/da.22228

WHO (2008). What is primary care mental health?. *Mental health in family medicine, 5*(1), 9–13.

Zagorscak, P., Heinrich, M., Bohn, J., Stein, J., & Knaevelsrud, C. (2020). How individuals change during internet-based interventions for depression: A randomized controlled trial comparing standardized and individualized feedback. *Brain and behavior, 10*(1). https://doi.org/10.1002/brb3.1484

Tables

Table 1

Table of Included Studies and Extracted Data for Synthesis

Citation	*N*	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
1. McDevitt-Petrovic et al., (2020)	253	Anxiety Disorders and Depression	Age: *M* = 38 (*SD* = 13.5) Gender: 152 (60%) females, 101 (40%) males Ethnicity: not reported	Low-intensity CBT	Individual	GMM	PHQ-9 GAD-7	Sessions 1 – 6 (6 time points)	11
2. Batterham et al., (2016)	417	Depression and Suicidal Ideation	Age: *M* = 40.6 (*SD* = 11.9) Gender: 323 (77%) females, 93 (23%) males Ethnicity: not reported	CBT	Digital	GMM	SIDAS CES-D	Assessment – 12-month follow-up (4 time points)	11
3. Moggia et al., (2020)	108	Depression	Age: *M* = 49.28 (*SD* = 10.98) Gender: 84 females (77.8%), 24 (22.2%) males Ethnicity: not reported	CBT Group-CBT Dilemma-focused therapy (DFT)	Combined: individual and group	GMM	CORE-SF	Assessment – post-treatment assessment (19 time points)	13

Citation	*N*	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
4. Zagorscak et al., (2020)	1089	Depression	Age: *M* = 45.7 (*SD* = 11.4) Gender: 714 (65.6%) females, 375 (34.4%) males Ethnicity: not reported	CBT	Digital	GMM	PHQ-9	Assessment – post-treatment assessment (8 time points)	13
5. Petersen et al., (2018)	626	Depression	Age: Control group - *M* = 51.2 (*SD* = 14.6); Intervention group - *M* = 50.3 (*SD* = 14.7) Gender: 473 females (75.56%), 153 (24.44%) males Ethnicity: not reported	Behavioural Activation	Individual	GMM	PHQ-9	Assessment – end of intervention (12 time points)	9
6. Lukaschek et al., (2019)	176	Panic Disorder or Panic Disorder with Agoraphobia	Age: *M* = 45.3 (*SD* = 13.8) Gender: 126 (71.6%) females, 50 (28.4%) males Ethnicity: not reported	CBT	Individual	GMM	OASIS	Assessment – session 10 (11 time points)	8
7. Santacana et al., (2016)	97	Panic Disorder or Panic Disorder with Agoraphobia	Age: *M* = 36.19 (*SD* = 9.23) Gender: 61 (62.9%) females, 36 (37.1%) males Ethnicity: not reported	Group CBT CBT (booster sessions)	Combined: individual and group	GMM	PDSS-SR	Assessment – follow up (13 time points)	9

Citation	*N*	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
8. Bray et al., (2016)	474	PTSD	Age: predominantly < 35 Gender: 386 (81.4%) females, 88 (18.6%) males Ethnicity: White = 219 (46.2%), Black = 118 (24.9%), Hispanic = 83 (17.5%), Other = 54 (11.4%)	CBT	Individual	GMM	PDS	Assessment – 12-month follow up (4 time points)	9
9. Lutz et al., (2014)*	326	Panic Disorder	Age: *M* = 37 (*SD* = 11.9) Gender: 210 (64.4%) females, 116 (35.6%) males Ethnicity: White = 286 (87.7%), African American = 16 (4.9%), Asian American = 18 (5.5%), Not specified = 6 (1.8%)	CBT	Individual	GMM	PDSS-SR	Session 1 – 5 (5 time points)	12
10. Koffman (2017)	251	Anxiety Disorders and Depression	Age: *M* = 30.6 (*SD* = 10.6) Gender: 158 (62.95%) females, 93 males (37.05%) Ethnicity: European Americans = 212 (84.46%), Minority backgrounds = 39 (15.54%)	CBT IPT	Individual	GMM	OQ-45	Session 1 – 6 (6 time points)	N/A
11. Flood et al., (2018)	1467	Anxiety Disorders and Depression	Age: *M* = 39.21 (*SD* = not reported) Gender: 874 females (59.6%), 593 (40.4%) Ethnicity: not reported	Group CBT	Group	GMM	DI-5 WHO-5	Session 1 – 5 (5 time points)	13

Citation	N	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
12. Hull et al., (2020)	10,718	Anxiety Disorders and Depression	Age: predominantly 26-35 Gender: 8456 (78.9%) females, 2262 (21.2%) males Ethnicity: not reported	CBT Third-wave CBT Psychodynamic / Relational	Digital	GMM	GAD-7 PHQ-9	Assessment – session 12 (5 time points)	12
13. Galovski et al., (2016)	69	PTSD	Age: not reported for full sample Gender: not reported Ethnicity: White = 34 (49.28%), Hispanic = 5 (7.25%), Not specified = 33 (47.83%)	Cognitive Processing Therapy (CPT)	Individual	LCGA	PDS BDI	Assessment – point of 'recovery' (session 4-18)	11
14. Wardenaar et al., (2014)	153	Depression	Age: *M* = 43.3 (*SD* = not reported) Gender: 105 (68.6%) females, 48 (31.38%) males Ethnicity: not reported	High-intensity CBT Low-intensity CBT	Individual	LCGA / GMM	BDI	Assessment – 12-month follow-up (over 52 data time points)	13
15. Lewis et al., (2012)*	173	Anxiety Disorders and Depression	Age: *M* = 27.94 (*SD* = 11.42) Gender: 115 (66.47%) females, 58 (33.53%) males Ethnicity: White = 160 (92.49%), Minority backgrounds = 13 (7.41%)	CBT	Individual	LCGM	BDI	Assessment – session 5 (6 time points)	N/A

Citation	N	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
16. Stulz & Lutz (2007)	1128	Anxiety Disorders and Depression	Age: M = 37.5 (SD = 9.5) Gender: 801 (71%) females, 327 (29%) males Ethnicity: White = 1015 (90%), Minority backgrounds = 113 (10%)	CBT Integrative approach	Individual	GMM	COMPASS	Assessment – session 17 (5 time points)	10
17. Lutz et al., (2020)	212	Anxiety Disorders and Depression	Age: M = 35.66 (SD = 12.72) Gender: 102 (48%) females, 110 (52%) males Ethnicity: not reported	CBT with "inter-personal and emotion-focused elements"	Individual	GMM	IIP-12	Assessment – post-treatment (4 time points)	9
18. Eftekhari et al., (2020)*	782	PTSD	Age: M = 46.9 (SD = 14.22) Gender: 359 (13.8%) females, 2246 (86.2%) males Ethnicity: White = 1873 (71.9%), Black = 498 (19.1%), Hispanic = 234 (9%)	Prolonged Exposure Therapy (PET)	Individual	GMM	PCL	Session 1 – 5 (5 time points)	7
19. Litz et al., (2019)	702	PTSD	Age: M = 32.88 (SD = 7.39) Gender: 74 (10.6%) females, 628 (89.4%) males Ethnicity: White = 393 (56%), Black or African American = 180 (26%), Native Hawaiian = 8 (1%), Asian = 8 (1%), American Indian = 16 (2%), Not specified = 97 (14%)	CBT	Individual	GMM	PSS-I	Assessment – 12-month follow-up (4 time points)	N/A

Citation	N	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
20. Bei et al., (2019)	148	Depression and Insomnia	Age: *M* = 46.56 (*SD* = 12.64) Gender: 108 (73%) females, 40 (27%) males Ethnicity: White = 105 (71%), Minority backgrounds = 53 (29%)	CBT	Digital	GMM	HRSD	Assessment – 2-year follow-up (15 time points)	12
21. Schuum et al., (2013)	207	PTSD	Age: *M* = 42.27 (*SD* = 14.42) Gender: 23 (11%) females, 184 (89%) males Ethnicity: White = 168 (81%), Minority backgrounds = 39 (19%)	Cognitive Processing Therapy (CPT)	Individual	GMM	PCL-S	Session 1 – 12 (12 time points)	11
22. Heckman et al., (2015)	105	Depression	Age: not reported for full sample Gender: 53 (50.48%) females, 50 (47.62%) males, 2 (1.9%) not specified Ethnicity: White = 23 (21.9%), Minority backgrounds = 82 (78.1%)	Coping Enhancement Group Teletherapy (CBT) Supportive–Expressive Group Teletherapy (Humanistic)	Group	GMM	GDS	Session 1 – 12 (12 time points)	7
23. Lutz et al., (2017)*	409	Depression	Age: *M* = 43.16 (*SD* = 11.10) Gender: 287 (70%) females, 122 (30%) males Ethnicity: not reported	CBT	Digital	GMM	PHQ-9	‘Screening’ – session 4 (4 time points)	12

Citation	*N*	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
24. Allan et al., (2017)	231	PTSD	Age: *M* = 45.70 (*SD* = 14.89) Gender: 13 (5.6%) females, 218 (94.4%) males Ethnicity: White = 117 (50.6%), Black = 109 (47.2%), Not specified = 5 (2.2%)	Behavioural Activation and Therapeutic Exposure (BA-TE)	Individual	GMM	PCL	Assessment – post-intervention (6 time points)	13
25. Thibodeau et al., (2015)	821	Depression	Age: *M* = 39.74 (*SD* = 10.61) Gender: 560 (68.2%) females, 261 (31.8%) males Ethnicity: not reported	CBT Supportive therapy Psychodynamic therapy	Individual	GMM	MADRS	Assessment – month-6 of treatment or follow-up (4 time points)	10
26. Crane et al., (2019)	293	Anxiety Disorders and Depression	Age: *M* = 48.96 (*SD* = 10.69) Gender = 293 (100%) females Ethnicity: not reported	Interpersonal Counselling	Individual	GMM	CES-D STAI / PROMIS	Assessment – 16-week follow-up (3 time points)	9
27. Sunderland et al., (2012)	663	Anxiety Disorders and Depression	Age: not reported for full sample Gender: 442 (66.67%) females, 221 (33.33%) males Ethnicity: not reported	CBT	Digital	GMM	PHQ-9 GAD-7 K10	Assessment – End of treatment (6 time points)	7

Citation	*N*	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
28. Nordberg et al., (2014)	147	Anxiety Disorders and Depression	Age: *M* = 39 *(SD* = 12) Gender: 104 (71%) females, 43 (29%) males Ethnicity: White = 147 (100%)	CBT Psychodynamic Therapy	Individual	GMM	TOP	Assessment – up to 4 months (3 time points)	13
29. Stein et al., (2012)	313	PTSD	Age: *M* = 33.60 (*SD* = 11.28) Gender: 313 (100%) females Ethnicity: not reported	Prolonged Exposure Therapy (PET) Cognitive Processing Therapy (CPT) – with and without written accounts Written Exposure Therapy (WET)	Individual	GMM (separate models for each study sample)	PSS / PDS	Assessment – post-treatment follow-up (9 time points)	11
30. Rubel et al., (2015)*	5484	Anxiety Disorders and Depression	Age: not reported Gender: 3384 (61.7%) females, 197 1903 (34.7%), males, (3.6%) not specified Ethnicity: European American = 2232 (40.7%), Asian American = 225 (4.1%), African American = 203 (3.7%), Latino / Hispanic = 165 (3%), Native American = 27 (0.5%), Mixed = 22 (0.4%), Other = 433 (7.9), Not specified = 2177 (39.7%)	Not Specified	Individual	GMM	BHM-20	Session 1 – 3 (3 time points)	6

Citation	*N*	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
31. Joesch et al., (2013)	482	Anxiety Disorders	Age: *M* = 43.43 (*SD* = 13.26) Gender: 339 (70.33%) females, 143 (29.66%) males Ethnicity: White = 273 (56.64%), Hispanic = 93 (19.29%), Black = 48 (9.96%), Other = 68 (14.11%)	CBT	Individual	GBTM	OASIS	Assessment – follow up (14 time points)	10
32.Stulz et al., (2007)*	192	Not specified	Age: *M* = 36.7 (*SD* = 10.8) Gender: 137 (71%) females, 55 (29%) males Ethnicity: not reported	CBT Psychodynamic Therapy Gestalt Therapy Cognitive Analytic Therapy Transactional Analysis Integrative approaches	Individual	GMM	CORE-SF	Session 1 – 6 (6 time points)	9
33. Malgaroli et al., (2020)	475	PTSD	Age: predominantly 36-35 Gender: 412 (86.7%) females, 63 (13.3%) males Ethnicity: not reported	CBT Third-wave CBT Psychodynamic Therapy	Digital	GMM	PCL-5	Assessment – end of treatment (5 time points)	11

Citation	N	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
34. Siddique et al., (2012)	267	Depression	Age: M = 29.3 (SD = 7.9) Gender: 267 (100%) females Ethnicity: Latina = 134 (50.2%), Black = 117 (43.8%), White = 16 (6%)	CBT Group CBT	Combined: individual and group	GMM	HDRS	Assessment – 12-month follow-up (10 time points)	12
35. Owen et al., (2015)	10,854	Not specified	Age: not reported Gender: 6706 (61.8%) females, 3424 (31.5%) males, 724 (6.7%) not specified Ethnicity: White = 5526 (67.7%), Asian / Pacific Islander = 651 (8%), African American = 450 (5.5%), Hispanic = 414 (5.1%), Native American = 67 (0.1%), Mixed = 186 (2.3%), Other = 868 (10.6%)	Counselling (type not specified)	Individual	GMM	BHM-20	Session 1 – 14 (14 time points)	7
36. Lutz et al., (2009)*	162	Depression	Age: not reported Gender: not reported Ethnicity: not reported	CBT Interpersonal Therapy (IPT)	Individual	GMM	BDI	Assessment – session 8 (3 time points)	9
37. Stulz et al., (2010)	504	Depression	Age: M = 42.8 (SD = 10.2) Gender: 323 (64%) females, 181 (36%) males Ethnicity: White = 464 (92%), Minority backgrounds = 40 (8%)	Cognitive behavioral analysis system of psychotherapy (CBASP)	Individual	GMM	HRSD	Assessment – session 12, end of treatment (9 time points)	11

Citation	*N*	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
38. Cuijpers et al., (2005)	226	Depression	Age: *M* = 36 (*SD* = 10) Gender: 138 (61%) females, 88 (39%) males Ethnicity: not reported	CBT Brief Therapy Clinician's choice	Individual	GMM	SCL	Assessment – 18-month follow-up (7 time points)	11
39. Saunders at al., (2019)	4394	Anxiety Disorders and Depression	Age: *M* = 38.49 (*SD* = 12.98) Gender: 2783 (63.24%) females, 1538 (35%) males, 73 (1.66%) not specified Ethnicity: White = 2963 (67.43%), Black = 253 (5.76%), Asian = 235 (5.35%), Mixed = 221 (5.03%), Other = 172 (3.91%), Not specified = 550 (12.52%)	CBT IPT Counselling	Individual	LCGA	PHQ-9 GAD-7	Assessment – session 12 (12 time points)	14
40. Phelps et al., (2017)	2686	PTSD	Age: *M* = 55.92 *(SD* = 10.54 Gender: 32 (1.2%) females, 2654 (98.8%) males Ethnicity: not reported	Cognitive-behavioural approaches	Individual	GMM	PCL	Assessment – 9-month follow-up (4 time points)	11
41. Smits et al., (2015)*	402	Not specified	Age: *M* = 28.27 (*SD* = 11.94) Gender: 261 (64.9%) females, 141 (35.01%) males Ethnicity: not reported	Humanistic / psychodynamic Therapy CBT Systemic Therapy	Individual	LCGA	OQ-45	Session 1 – 5 (5 time points)	10

Citation	N	CMD(s) of sample	Demographics	Intervention(s)	Medium of Delivery	GCM Approach	Measure(s)	Time Points	GRoLTS Score
42. Murphy & Smith (2018)	960	PTSD	Age: M = 42.99 (SD = 10.61) Gender: not reported Ethnicity: not reported	Group CBT	Group	LCGA	IES-R	Assessment – 12-month follow-up (5 time points)	13
43. Frankfurt et al., (2019)	508	PTSD	Age: M = 36.85 (SD = 10.01) Gender: 204 (40%) females, 304 (60%) males Ethnicity: White = 321 (63%), Minority backgrounds = 187 (27%)	Written Exposure Therapy (WET)	Individual	LCGA	GSI	Assessment – 6-month follow-up (3 time points)	17
44. Batterham et al., (2017)	1149	Depression and Insomnia	Age: M = 43 (SD = 12) Gender: 850 (74%) females, 299 (26%) males Ethnicity: not reported	Cognitive-behavioural online insomnia intervention (SHUTi)	Digital	GMM	PHQ-9NS (excludes sleep item)	Assessment – 18-month follow-up (9 time points)	4
45. Fletcher et al., (2017)	439	PTSD	Age: M = 36.46 (SD = 10.83) Gender: 377 (85.8%) females, 62 (14.12%) males Ethnicity: 439 (100%) White	CBT Psychodynamic Therapy	Individual	LCGA	HTQ	Assessment – 18-month follow-up (4 time points)	13

Note. Studies denotated with a '*' indicate those considered explorations of 'early change'. GMM = Growth Mixture Modelling; LCGA = Latent Class Growth Analysis; GBTM = Group Based Trajectory Modelling.

Table 2

Table of Trajectory Class Frequencies per Included Study

		Trajectory Class Frequency Counts (%)						
		Responders						
Citation	Symptom or Domain of Interest	Unspecified Improvement	Delayed Responders	Rapid Responders	Non-responders	Deteriorators	Other	Total
1. McDevitt-Petrovic et al., (2020)a	Anxiety	1 (82.80%)	-	-	1 (17.20%)	-	-	2 (*n* = 250)
McDevitt-Petrovic et al., (2020)b	Depression	1 (87.25%)	-	-	1 (12.75%)	-	-	2 (*n* = 251)
2. Batterham et al., (2019)a	Suicidal Ideation	1 (83%)	-	-	1 (17%)	-	-	2 (n = 418)
Batterham et al., (2019)b	Depression	1 (81%)	-	-	1 (19%)	-	-	2 (*n* = 418)
3. Moggia et al., (2020)	Depression	-	-	-	1 (24.1%)	-	2 OT1 (58.3%) OT2 (17.6%)	3 (*n* = 108)
4. Zagorscak et al., (2019)	Depression	-	1 (37.5%)	1 (62.5%)	-	-	-	2 (*n* = 1089)

		Trajectory Class Frequency Counts (%)						
		Responders						
Citation	Symptom or Domain of Interest	Unspecified Improvement	Delayed Responders	Rapid Responders	Non-responders	Deteriorators	Other	Total
5. Petersen et al., (2018)	Depression	-	1 (39.5%)	1 (60.5%)	-	-	-	2 (*n* = 626)
6. Lukaschek et al., (2019)	Anxiety (Panic)	-	1 (33%)	1 (50.5%)	1 (16.5%)	-	-	3 (*n* = 176)
7. Santacana et al., (2016)	Anxiety (Panic)	1 (80.3%)	-	-	1 (19.7%)	-	-	2 (*n* = 97)
8. Bray et al., (2016)	PTSD	1 (18%)	-	-	1 (82%)	-	-	2 (*n* = 474)
9. Lutz et al., (2014)	Anxiety (Panic)	-	1 (58%)	1 (20.2%)	1 (17.2%)	1 (4.6%)	-	4 (*n* = 326)
10. Koffman (2017)	Symptom Distress	1 (100%)	-	-	-	-	-	1 (*n* = 251)
11. Flood et al., (2018)a	Psychological Distress	2 (50.7%; 35.99%)	-	-	1 (13.29%)	-	-	3 (*n* = 1467)
Flood et al., (2018)b	Mental Wellbeing	2 (14.25%; 31.70%)	-	-	1 (54.06%)	-	-	3 (*n* = 1467)

		Trajectory Class Frequency Counts (%)						
		Responders						
Citation	Symptom or Domain of Interest	Unspecified Improvement	Delayed Responders	Rapid Responders	Non-responders	Deteriorators	Other	Total
12. Hull et al., (2020)	Depression and Anxiety	3 (23.7%; 20%; 16.9%)	-	1 (7%)	2 (22.6%; 9.8%)	-	-	6 (*n* = 10,718)
13. Galovski et al., (2016)a	PTSD	1 (44.9%)	1 (7.2%)	1 (47.8%)	-	-	-	3 (*n* = 69)
Galovski et al., (2016)b	Depression	1 (46.4%)	1 (47.8%)	1 (5.8%)	-	-	-	3 (*n* = 69)
14. Wardenaar et al., (2014)	Depression	-	1 (9.8%)	1 (40.2%)	1 (33%)	-	1 OT3 (17%)	4 (*n* = 153)
15. Lewis et al., (2012)	Depression and Anxiety	1 (100%)	-	-	-	-	-	1 (*n* = 173)
16. Stulz & Lutz (2007)	Mental Wellbeing, Symptom Severity & Functioning	2 (63%; 20%)	-	1 (17%)	-	-	-	3 (*n* = 1128)
17. Lutz et al., (2020)	Interpersonal Problems	-	1 (68.4%)	1 (5.66%)	-	1 (25.94%)	-	3 (*n* = 212)
18. Eftekhari et al., (2020)	PTSD	1 (92.3%)	-	1 (7.2%)	-	1 (0.7%)	-	3 (*n* = 782)

		Trajectory Class Frequency Counts (%)						
		Responders						
Citation	Symptom or Domain of Interest	Unspecified Improvement	Delayed Responders	Rapid Responders	Non-responders	Deteriorators	Other	Total
19. Litz et al., (2019)	PTSD	1 (100%)	-	-	-	-	-	1 (n = 702)
20. Bei et al., (2019)	Depression	1 (17.6%)	1 (68.9%)	1 (13.5%)	-	-	-	3 (n = 148)
21. Schumm et al., (2013)	PTSD	-	1 (24.3%)	1 (57.1%)	1 (18.6%)	-	-	3 (n = 207)
22. Heckman et al., (2015)	Depression	-	1 (16.2%)	1 (30.5%)	1 (53.3%)	-	-	3 (n = 105)
23. Lutz et al., (2017)	Depression	-	-	-	-	1 (16%)	2 OT4 (38.6%) OT5 (45.2%)	3 (n = 409)
24. Allan et al., (2017)	PTSD	1 (15%)	-	1 (3%)	1 (82%)	-	-	3 (n = 231)
25. Thibodeau et al., (2014)	Depression	2 (42%; 31%)	-	1 (11%)	1 (16%)	-	-	4 (n = 821)

		Trajectory Class Frequency Counts (%)						
		Responders			Non-responders	Deteriorators	Other	Total
Citation	Symptom or Domain of Interest	Unspecified Improvement	Delayed Responders	Rapid Responders				
26. Crane et al., (2019)a	Depression	1 (7%)	-	-	2 (78%; 15%)	-	-	3 (*n* = 293)
Crane et al., (2019)b	Anxiety	1 (18%)	-	-	1 (73%)	1 (9%)	-	3 (*n* = 215)
27. Sunderland et al., (2012)a	Anxiety	1 (80%)	-	-	1 (20%)	-	-	2 (*n* = 361)
Sunderland et al., (2012)b	Depression	1 (75%)	-	-	1 (25%)	-	-	2 (*n* = 302)
28. Nordberg et al., (2014)	Depression	-	-	1 (14%)	2 (58%; 28%)	-	-	3 (*n* = 147)
29. Stein et al., (2012)a	PTSD	1 (96.1%)	-	-	1 (3.69%)	-	-	2 (*n* = 163)
Stein et al., (2012)b	PTSD	1 (76.67%)	-	-	1 (23.33%)	-	-	2 (*n* = 150)
Stein et al., (2012)c	PTSD	1 (87%)	-	-	1 (13%)	-	-	2 (*n* = 313)

		Trajectory Class Frequency Counts (%)						
		Responders						
Citation	Symptom or Domain of Interest	Unspecified Improvement	Delayed Responders	Rapid Responders	Non-responders	Deteriorators	Other	Total
30. Rubel et al., (2015)	Global Mental Health (Wellbeing, Symptoms & Functioning)	-	2 (27.7%; 62.7%)	1 (7.2%)	-	1 (2.4%)	-	4 (*n* = 5484)
31. Joesch et al., (2013)	Anxiety	1 (40.1%)	1 (3.7%)	1 (27.6%)	1 (3.7%)	-	-	4 (*n* = 482)
32. Stulz et al., (2007)	Psychological Distress	1 (29.17%)	-	1 (12.50%)	3 (13.54%; 25%; 19.79%)	-	-	5 (*n* = 192)
33. Malgaroli et al., (2020)	PTSD	2 (41.4%; 41.4%)	-	1 (4.3%)	1 (12.9%)	-	-	4 (*n* = 475)
34. Siddique et al., (2012)	Depression	2 (35%; 65%)	-	-	-	-	-	2 (*n* = 267)
35. Owen et al., (2015)	Global Mental Health (Wellbeing, Symptoms & Functioning)	1 (19.3%)	-	-	-	-	2 OT6 (75.3%) OT7 (5.4%)	3 (*n* = 10,854)
36. Lutz et al., (2009)	Depression	2 (19.8%; 19.1%)	-	1 (61.1%)	-	-	-	3 (*n* = 162)
37. Stulz et al., (2010)	Depression	3 (21.43%; 58.33%; 20.24%)	-	-	-	-	-	3 (*n* = 504)

		Trajectory Class Frequency Counts (%)						
		Responders						
Citation	Symptom or Domain of Interest	Unspecified Improvement	Delayed Responders	Rapid Responders	Non-responders	Deteriorators	Other	Total
38. Cuijpers et al., (2005)	Depression	1 (26%)	-	2 (31%; 33%)	1 (10%)	-	-	4 (*n* = 226)
39. Saunders et al., (2019)a	Depression	-	1 (13.5%)	2 (44.5%; 14%)	1 (28%)	-	-	4 (*n* = 4394)
Saunders et al., (2019)b	Anxiety	-	1 (11.6%)	2 (28.6%; 17.4%)	2 (21.4%; 21.1%)	-	-	5 (*n* = 4394)
40. Phelps et al., (2017)	PTSD	3 (3%; 49.9%; 6.7%)	-	1 (7.9%)	1 (32.5%)	-	-	5 (*n* = 2686)
41. Smits et al., (2015)	Psychological Distress	3 (30.5%; 20.4%; 41.3%)	-	-	1 (8%)	-	-	4 (*n* = 402)
42. Murphy & Smith (2018)	PTSD	3 (2.7%; 22.9%; 45.7%)	-	-	1 (27.5%)	-	1 OT8 (1.2%)	5 (*n* = 960)
43. Frankfurt et al., (2019)	PTSD	2 (60%; 5%)	-	-	1 (25%)	1 (10%)	-	4 (*n* = 508)
44. Batterham et al., (2017)	Depression	1 (95%)	-	-	1 (5%)	-	-	2 (*n* = 1149)

		Trajectory Class Frequency Counts (%)						
		Responders			Non-responders	Deteriorators	Other	Total
Citation	Symptom or Domain of Interest	Unspecified Improvement	Delayed Responders	Rapid Responders				
45. Fletcher et al., (2017)	PTSD	2 (26.22%; 33.03%)	-	1 (15.71%)	1 (15.03%)	-	-	4 (n = 439)
Totals		59	16	30	43	7	8	163

Note. ‘a/b/c’ denotes separate growth curve models fitted for different measures of the same study. OT1 = Improvement during group therapy; OT2 = “Improvement during 1-1 therapy”; OT3 = “Remission, Recurrence”; OT4 = “Early response after registration”; OT5 = “Early response after screening”; OT6 = “Early & Late Change”; OT7 = “Worse Before Better”; OT8 = “Rapid response, Remit".

Table 3

Table of Trajectory Class Frequencies by Intervention Type, CMD and Delivery Medium

Trajectory Class		Intervention Type				
		Cognitive-Behavioural (CBT)	Mixed Provisions	Counselling	Not Specified	Total
Responders	Unspecified Improvement	37	19	3	0	59
	Delayed Responder	11	3	0	2	16
	Rapid Responder	14	15	0	1	30
Non-Responders		24	16	3	0	43
Deteriorators		5	0	1	1	7
Other		6	0	2	0	8
Total		97	53	9	4	163

Trajectory Class		CMD					
		Anxiety Disorders	Depression Disorders	Anxiety & Depression Disorders	PTSD	Not Specified	Total
Responders	Unspecified Improvement	2	14	17	21	5	59
	Delayed Responder	3	5	5	3	0	16

	Rapid Responder	3	9	9	8	1	30
Non-Responders		4	8	16	11	4	43
Deteriorators		1	1	3	2	0	7
Other		0	5	0	1	2	8
Totals		13	42	50	46	12	163

Trajectory Class		Delivery Medium				
		Individual	Group	Digital	Combined: Individual & Group	Total
	Unspecified Improvement	38	7	11	3	59
Responders	Delayed Responder	13	1	2	0	16
	Rapid Responder	25	1	4	0	30
Non-Responders		30	4	8	1	43
Deteriorators		6	0	1	0	7
Other		3	1	2	2	8
Totals		115	14	28	6	163

Table 4

Table of Patient Characteristics Associated with Class Membership

Predictor Type	Trajectory Class					
	Responders			Non-Responders	Deteriorators	Other
	Unspecified Improvement	Delayed Responders	Rapid Responders			
Baseline severity	Greater behavioural health (BHM-20)[35] Greater psychological distress (OQ-45)[41]	Greater PTSD symptoms (PDS scores)[13] Greater depression symptoms (BDI)[13b] Lower PTSD symptoms rated by patient and therapist (PCL)[21] Greater depression symptoms (PHQ-9)[39a] Lower anxiety symptoms (GAD-7)[39b]	Lower anxiety symptoms (OASIS)[6] Lower depression symptoms (HRSD)[20] Lower PTSD symptoms rated by patient and therapist (PCL)[21] Greater psychological impairment (lower GMH score)[30] Lower depression symptoms (PHQ-9)[39a] Greater depression symptoms (PHQ-9)[39a]	Greater anxiety symptoms (GAD-7)[2] Greater depression symptoms (CORE-SF)[3] Greater PTSD symptoms rated by patient and therapist (PCL)[21] Greater depression symptoms (MADRS)[25] Lower depression symptoms (CES-D)[26a] Lower depression symptoms (CES-D)[26b] Greater distress (K10) and anxiety symptoms (GAD-7)[27a] Greater distress (K10) and depression symptoms (PHQ-9)[27b]	Greater depression symptoms (HRSD)[23]	*Group Therapy Improvers:* Lower depression symptoms (BDI-II)[3] Lower psychological distress (CORE-SF; distress subscale)[3] *Individual Therapy Improvers:* Greater depression symptoms (BDI-II)[3] *Early Response after Registration:*

Baseline severity cont.			Lower anxiety symptoms (GAD-7)[39b]	Lower depression symptoms (TOP – depression subscale)[28] Greater depression symptoms (BDI)[32] Greater depression symptoms (SCL)[38] Greater psychological distress (OQ-45)[41] Greater depression symptoms (PHQ-9)[44]		Higher depression symptoms (PHQ-9)[23] Higher depression symptoms (HRSD)[23] *Early and Late Change:* Greater behavioural health (BHM-20)[35]
Functioning	-	Greater functioning – lower scores (WSAS)[9]	Poorer functioning – greater scores (WSAS)[9] Greater functioning – lower scores (WSAS)[39a] Greater functioning – lower scores (WSAS)[39b]	Poorer functioning – greater scores (WSAS)[9] Greater impairment (GAF)[25] Greater impairment (WHODAS 2.0)[27a] Greater impairment (WHODAS 2.0)[27b] Poorer sexual functioning (TOP – sexual functioning subscale)[28] Poorer functioning (WSAS)[39b]	Greater functioning – lower scores (WSAS)[9]	*Group Therapy Improvers:* Higher functioning (GAF)[3]

Demographics	Younger (17-34)[8] Single (Marital Status)[8] Junior rank of army employment[8] Greater education attainment (college degree or higher)[12] Female[12] Younger[26a] Younger[26b] Married[31] Male[45]	Female[6] Younger[21] Unemployed[31]	Higher education attainment (at least a bachelor's degree)[12] Female[12] Married[24] Married[31] Older[32]	Single (Marital Status)[2] Unemployment1 'Not currently in labor force'[2] Urban living[6] Older (>35 years)[8] BAME ethnicity[21] Single (Marital Status)[24] Unemployed[31] Younger[44]	-	*Remission and Recurrence:* Married[14]
Diagnoses	Less likely to have 'probable PTSD diagnosis'[43]	MDD diagnosis[4] MDD diagnosis[6] Greater number of anxiety disorder diagnoses[31] MDD diagnosis[31]	GAD or 'other' diagnosis[39a] Phobic Anxiety or Panic diagnosis [39a] Other diagnosis[39a] No OCD diagnosis[39b] Other diagnosis[39b]	Comorbid mental health problem[1] Mixed depressive and anxiety disorder diagnosis[1] MDD diagnosis[6] Dysthymia diagnosis[14] Personality Disorder diagnosis[25] MDD diagnosis[29] MDD diagnosis[31]	-	*Remission and Recurrence:* Dysthymia diagnosis[14]

Diagnoses cont.				Greater number of anxiety disorder diagnoses[31] Dysthymic disorder diagnosis[38] Anxiety disorder diagnosis[38]		
Symptomology, onset and course of CMD	Lower guilt symptoms (CAPS IV – Guilt)[40] Lower re-experiencing and avoidance symptoms (HTQ)[45]	Previous depression episodes[14] Previous insomnia episodes[20] Previous depression episodes[20] Insomnia onset after current depression episode[20]	Lower re-experiencing and avoidance symptoms (HTQ)[45]	Onset of greater than 5 years[1] Lower age of onset[14] Fewer somatic symptoms (CIDI)[14] Greater hyperarousal symptoms (PSS / PDS – hyperarousal subscale)[29c] Greater guilt symptoms (CAPS IV – Guilt)[40] Greater re-experiencing and avoidance symptoms (HTQ)[45]	-	-

Social factors	Lower perceived burdensomeness (Interpersonal Needs Questionnaire; INQ)[2] Greater social wellbeing (QOL – social wellbeing subscale)[26a] Greater social wellbeing (QOL – social wellbeing subscale)[26b] Greater social support (Medical Outcomes Study Checklist)[31] Co-habiting with partner[31] Greater perceived social support (DRRI)[43] Greater perceived social support during and after trauma (CSS)[45]	Lower perceived social support (BSSS)[4]	Greater perceived social support (BSSS)[4] Greater social support (Medical Outcomes Study Checklist)[31] Co-habiting with partner[31] Greater perceived social support during and after trauma (CSS)[45]	Greater perceived burdensomeness (Interpersonal Needs Questionnaire; INQ)[2a] Greater Thwarted Belongingness (Interpersonal Needs Questionnaire; INQ)[2a] Greater perceived burdensomeness (Interpersonal Needs Questionnaire; INQ)[2b] Greater social conflict (TOP - social conflict subscale)[28] Lower perceived social support during and after trauma (CSS)[45]	-	*Group Therapy Improvers*: Lower perceived social isolation ('Self-others discrepancy)[3]

Personality factors	Lower emotion focused coping (CSQ)[45]	-	Lower emotion focused coping (CSQ)[45]	Greater neuroticism (NEO-FFI)[14] Lower extraversion (NEO-FFI)[14] Lower sense of mastery (Mastery Scale)[14] Greater dysfunctional attitudes (DAS)[25] Greater emotion focused coping (CISS)[25] Greater introversion (D5D)[25] Greater neuroticism (D5D)[25] Greater emotion / detachment focused coping (CSQ)[45]	-	-
Trauma characteristics	Less combat exposure[8] Less combat exposure[43] Longer time since abuse[45]	-	-	Greater combat exposure[8] Greater combat exposure[21] Combat role[42] Shorter time since abuse[45]	-	*Group Therapy Improvers*: Greater self-esteem ('Self-ideal discrepancy)[3] Lower dilemmatic construction of the self (PICID)[3]

Trauma characteristics cont.						*Individual Therapy Improvers:* Greater dilemmatic construction of the self (PICID)[3]
Risk to self and others	-	Prior suicide attempts(s)[5]	Increased hostility towards others (TOP – hostility subscale)[28]	Some risk of suicide (therapist assessed)[1] Neglect of self and others[1] Priors suicide attempt(s)[14]	-	-
Co-morbid mental health symptoms	Greater anxiety symptoms (GAD-7)[2] Greater Insomnia (Insomnia Severity Index; ISI)[2] Greater traits of avoidant personality disorder (SNAP)[13a] Lower anxiety symptoms (BAI)[24]	Lower depression symptoms (HAM-D)[9] Lower anxiety symptoms (HAM-A)[9] Lower separation anxiety in adulthood (ASAC)[9] Greater depression symptoms (BDI-II)[13a]	Greater separation anxiety in adulthood (ASAC)[9] Lower insomnia symptoms (ISI)[20] Healthier insomnia-related constructs – lower sleep effort (GSES) and fewer dysfunctional sleep beliefs (DBAS)[20] Lower depression symptoms (BDI-II)[21]	Severe insomnia symptoms (Insomnia Severity Index; (ISI)[2] Greater separation anxiety in adulthood (ASAC)[9] Greater depression symptoms (BDI-II)[21] Greater anxiety symptoms (BAI)[24] Greater suicidal ideation (TOP – suicidal ideation subscale)[28]	Lower depression symptoms (HAM-D)[9] Lower anxiety symptoms (HAM-A)[9] Lower separation anxiety in adulthood (ASAC)[9]	-

Co-morbid mental health symptoms cont.		Lower depression symptoms (BDI-II)[21] Lower anxiety symptoms (GAD-7)[39a] Lower phobia and avoidance symptoms (IAPT phobia scale)[39b]	Lower anxiety symptoms (BAI)[24] Lower anxiety symptoms (GAD-7)[29a] Lower phobia and avoidance symptoms (IAPT phobia scale)[39a] Lower depression symptoms (PHQ-9)[39b] Lower phobia and avoidance symptoms (IAPT phobia scale)[39b] Lower symptoms of depression (HADS)[40]	Greater depression symptoms (HADS)[40] Greater depression symptoms (PHQ-9)[42] Greater anxiety symptoms (GAD-7)[42]		
Physical health	Received Chemotherapy[26a]	More physical health conditions (documented in records)[5] Lower health-related quality of life (EQ-5D)[5] Greater number of major chronic medical conditions[31]	Fewer number of years living with HIV[22]	Fibromyalgia diagnosis[3] Greater number of chronic diseases[14] Greater number of years living with HIV[22] Received Chemotherapy[26a]	Poorer physical health (SF-12 Physical Health subscale)[23]	*Early Response after Screening:* Poorer mental health (F-12 Mental Health subscale)[23]

Physical health cont.				Greater number of major chronic conditions[31]		
Treatment factors	Greater weekly written communication (engagement)[12] More likely to drop-out[13a] Greater therapeutic alliance (WAI)[33] First experience of therapy[33] More engaged[33]	Greater initial expectations of treatment[4] Less likely to have a ‘High-intensity’ treatment[14] Greater therapist-patient nonverbal synchrony[17] Lower therapy group cohesion until session four (MGES cohesion subscale)[22] Greater perceived group member similarity (Group Member Similarity Scale) [22]	Greater treatment adherence[12] More likely to have a CBT therapist[12] Greater treatment completion[12] More engaged[12] More likely to drop-out[13a] Past treatment[16] Lower therapist-patient nonverbal synchrony[17] Greater therapy group cohesion (MGES cohesion subscale)[22] Greater therapeutic alliance (WAI)[33] First experience of therapy[33] More engaged[33]	Past treatment[1] Concurrent treatment[1] Use of anti-depressant[1] Less engaged[12] Poorer therapy group cohesion (MGES cohesion subscale)[22] Provided written accounts of traumas during therapy[29c] Poorer therapeutic alliance (WAI)[33] Limited comfort in using internet as treatment medium[44]	Greater therapist-patient nonverbal synchrony[17]	-

Note. Number in superscript corresponds to the citations of Table 1. ‘a/b/c’ denotes separate growth curve models fitted for different measures of the same study.

Part 2: Empirical Paper

Do symptom change trajectories differ between Cognitive Behavioural Therapy (CBT) and Counselling? An application of growth curve modelling using a large IAPT dataset

Abstract

Background: Psychological interventions are often effective in reducing symptoms of common mental disorders (CMD). However, there are notable within- and between-person differences in how symptoms change over time. To investigate such differences, some studies have used growth curve modelling (GCM) to identify different trajectory classes, allowing for the earlier recognition of distinct patient outcomes through routine outcome monitoring (ROM). These are often associated with various patient characteristics, informing the likelihood of a patient's trajectory during the early phases of treatment, but less is known about how these differ between specific treatments.

Aims: The study aimed to investigate whether the trajectory classes identified differed between individuals receiving Cognitive Behavioural Therapy (CBT) and Counselling, using data from the Improving Access to Psychological Therapies (IAPT) programme.

Method: The current study employed growth mixture modelling (GMM) to identify trajectory classes of depression and anxiety symptoms among a propensity-score matched sample (N = 10268) to control for pre-treatment confounding factors. Multinomial regression models were fitted to determine the associations between the classes and the two intervention types.

Results: Four classes were identified for both depression and anxiety symptom change: (1) Rapid Responders (2) Delayed Responders (3) Low Severity Small Improvers and (4) Non-responders. Participants who received Counselling were less likely to be Rapid Responders according to change in their depression symptoms and less likely to be Delayed Responders with regards to anxiety symptoms.

Conclusions: Through recognition of these classes during the delivery of CBT or Counselling, clinicians may be better placed to make treatment decisions that optimise patient outcomes.

Introduction

The guidance of the National Institute for Health and Care Excellence (NICE) recommends evidence-based psychological interventions for individuals experiencing symptoms of common mental disorders (CMD). These include depression, post-traumatic stress disorder (PTSD) and anxiety disorders, which are estimated to affect approximately a quarter of individuals at some point over the lifespan (NICE, 2011; Steel et al., 2014). To accommodate variations in the severity of CMD symptoms and their associated functional impairment, the NICE guidance follows stepped care principles, such that low-intensity interventions might be recommended for less severe presentations and high-intensity ones are recommended for moderate-to-severe presentations (NICE, 2011). In England, the national Improving Access to Psychological Therapies (IAPT) programme provides psychological therapies in line with NICE guidance to over a million adults each year (Clark, 2018; NHS Digital, 2020). The majority of IAPT service users receive cognitive-behavioural interventions, but for those with depression, most IAPT services also offer Counselling, Interpersonal Psychotherapy (IPT), Behavioural Couples Therapy (BCT) and Dynamic Interpersonal therapy (DIT). In contrast to Counselling and IPT, DIT has not been included in the most recent NICE guidelines for depression (NICE, 2009). However, it has a growing evidence base, with randomised controlled trials (RCT) indicating efficacy (Fonagy et al., 2020).

Alongside stepped care and evidence-based service delivery, IAPT services are also renowned for the key feature of routine outcome monitoring (ROM), whereby patients complete a series of symptom measures on a sessional basis regardless of their diagnosis, making up the IAPT minimum dataset (MDS; The National Collaborating Centre for Mental Health, 2020). This is largely comprised of the Patient Health Questionnaire-9 (PHQ-9) and the Generalised

Anxiety Disorder Scale-7 (GAD-7), measuring depression and anxiety symptoms, respectively (Kroenke, Spitzer, & Williams, 2001; Spitzer et al., 2006). Through plotting scores on the MDS measures over the course of treatment, individual changes in symptoms are determined and can be used to track progress and determine outcomes at the end of treatment, benefiting patients and therapists alike. Such outcomes are aggregated for each service and used to evaluate the performance of all IAPT services nationally every month, judging services against key performance indicators set by NHS England (Clark et al., 2018). The datasets have also been used by researchers to explore the real-world effectiveness of IAPT delivered psychological interventions, outside of the optimal conditions of RCTs. A recent meta-analysis of these studies found that on average, symptoms of anxiety and depression reduced at post-treatment with large effect sizes. However, high drop-out and relapse rates were highlighted as areas for future improvement (Wakefield et al., 2020). Comparisons of intervention effectiveness have also been made, with one study highlighting the near equivalent outcomes of CBT and Counselling, on average (Pybis et al., 2017).

It is important to note that not all patients experience symptomatic improvements over the course of treatment in IAPT. Outcomes are influenced by an interaction between patient, therapist, intervention and service-level variables, which can culminate in a limited response to therapy for some, and clinical deterioration for others (Amati et al., 2017; Clark et al., 2018). To explore this statistically, growth curve modelling (GCM) can be used to assess between- and within-person longitudinal changes in symptom measure scores and subtypes of GCM, such as growth mixture modelling (GMM) or latent class growth analysis (LGCA), can identify distinct classes of symptom change, for which membership can be predicted using regression analyses. These predictions can be made either before treatment has started or after initial sessions have taken place. Identifying distinct trajectory classes and predicting trajectory class membership together is known as a 'two-stage approach'. If applied to a clinical context, findings of such analyses can aid treatment planning through predicting an individual's response to an intervention. Research has found that the clinical judgment of therapists is often biased, made evident by overestimates of patient improvement rates and underestimates of deterioration. Therefore, regression models in line with the two-stage approach may drastically enhance patient outcomes by limiting biased decision making (Walfish et al., 2012). Moreover, those who are seemingly unlikely to respond to psychological intervention can be supported to access alternatives, for example therapies offered by secondary-care or 'combination treatment' (therapy with the addition of psychotropic medication). Although examples are limited in the relevant research literature, Kazdin (2007) has also outlined the potential of GCM to uncover differences in the underpinning mechanisms of change of psychological interventions, as reflected by observable differences in trajectory classes.

The systematic review that informed the current study (part one of this thesis) aimed to compare symptom class trajectories between differing psychological interventions for CMDs, as well as their predictors. Across 45 studies, three main classes of CMD symptom change were identified: "Responders", "Non-responders" and "Deteriorators", with the former subclassified as either "Unspecified Improvers", "Delayed Responders" or "Rapid Responders". Eight additional classes were also identified that did not fit the aforementioned main classes due to unique shapes and form. These are described in full in part one of the thesis. Almost all studies included in the review identified at least one "Responder" class and by and large, the majority of participants belonging to it. Meanwhile, fewer participants were found to belong to the "Non-responder" or "Deteriorator" trajectory classes. The same pattern was observed when the trajectories were grouped by CMD, further implying that classes could be universal and generally not impacted by the diagnosis a person has. The predictors of trajectory classes were widespread, although few were consistent across all included studies. Several patient-level characteristics were identified as potentially helpful for predicting therapy response, including baseline symptom severity, mental or physical health comorbidities and social support. One notable study by Saunders and colleagues (2019), used LGCA on a large IAPT data set, producing a four-class model for PHQ-9 and a five-class model for GAD-7 scores over the course of high-intensity psychological therapy. However, as with the other studies included in the review, the type of therapy (e.g., CBT vs Counselling) was not explored as a predictor of class membership despite a range of therapies received by participants in the sample, and average (mean) growth curves distinguished on the basis of these therapies. As such, the study was unable to provide information to support clinical decisions about what type of therapy may be most likely to benefit any individual patient, or what type of symptom change a patient may be most likely to experience with any given type of therapy. This is fundamental to answering questions of "what works for whom?", which could support personalised treatment recommendations and better-informed collaborative treatment decisions between clinician and patient (Cohen & DeRubeis, 2018; DeRubeis et al., 2014; Roth & Fonagy, 2005).

The current study has several aims. First, to explore whether there are differences in the mean rate of depression and anxiety symptom change for the two most commonly received high-intensity therapies in IAPT (CBT and Counselling). Second, to identify trajectory classes of symptom change and describe their shape and form. Finally, to ascertain associations between class membership and the type of treatment received (CBT or Counselling).

Method

Setting

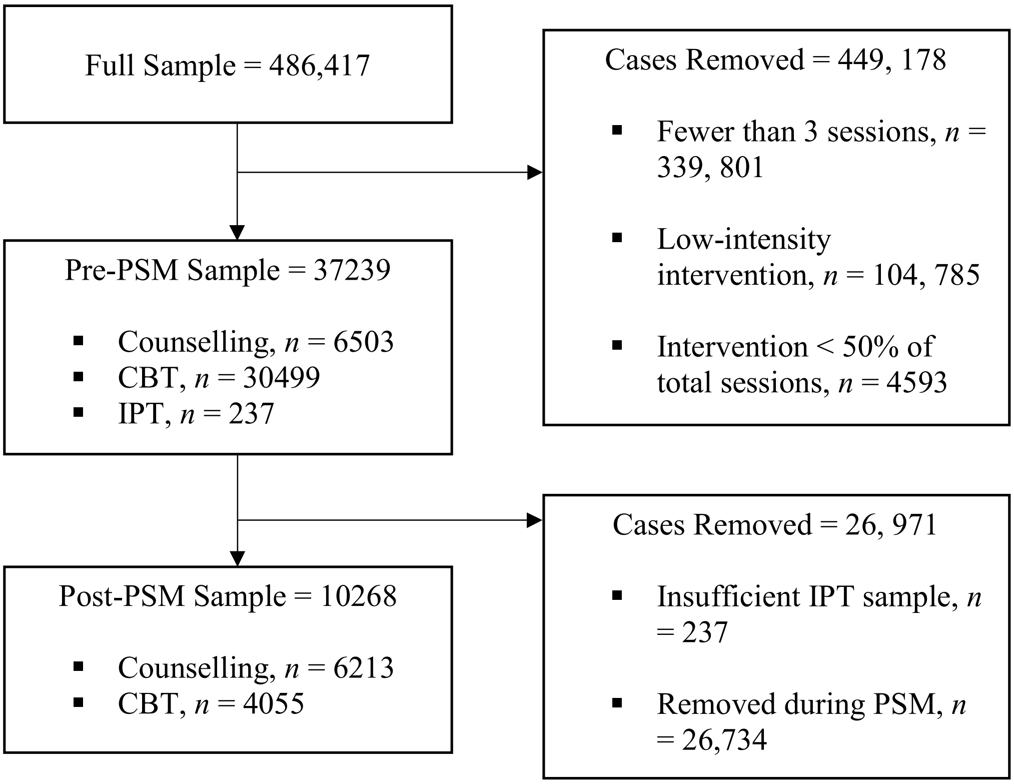
The data used to address the aims of the study was taken from eight IAPT services, across four NHS Trusts, based in North, Central and East London (NCEL). All details of the dataset are provided by Saunders and colleagues (2020). Each service provides psychological therapies to adults with CMDs that have either been referred by a health professional (most commonly a General Practitioner) or have self-referred. The services are commissioned by area-specific clinical commissioning groups (CCG) and receive differing allocations of funds. Given the diversity observed in London, these services are intended to provide care to individuals from various communities and backgrounds (Greater London Authority, 2021). According to the 2011 census, 44.9% of Londoners are White British and 37% were born outside of the United Kingdom, highlighting the requirement of these services to be culturally sensitive (Office for National Statistics, 2011). To facilitate this, guidance has been issued specifically for IAPT services by organisations such as the British Association of Behavioural and Cognitive Psychotherapies (BABCP; Beck et al., 2019).

Participants

The sample was comprised of individuals receiving either CBT of Counselling, both high-intensity psychological therapies, from one of the eight services. Because a comparably small proportion of the sample had sessions of IPT, these participants were removed despite an initial plan to include the intervention as part of subsequent analyses. Participants were excluded from the sample if they did not attend a minimum of three sessions, so that appropriate analyses could be conducted. To ensure a good standard of intervention fidelity, participants were also removed if CBT or Counselling did not comprise at least half of their total sessions. Figure 1 provides a flow-chart of the procedure used to define the study sample.

Figure 1

A CONSORT Style Flow-chart to Represent the Sample Size before and after the Application of Exclusion Criteria and PSM (Schulz et al., 2010)



Note. PSM = Propensity score matching.

Interventions

In IAPT services, a particular type of therapy is selected to treat a corresponding CMD or ‘presenting problem’, informed by NICE guidance. A summary of the matching of treatments to CMD in IAPT is provided by Clark (2018) and presented in Appendix A. To deliver these treatments, high-intensity therapists are trained in line with frameworks that define the competencies required to support individuals with depression and anxiety disorders (Roth & Pilling, 2007). Over time, the scope of these frameworks has expanded beyond CBT to describe competencies relevant to Counselling approaches, mirroring the widening access of these interventions (Hill, 2010; Roth et al., 2009). In recent years, IAPT services have begun offering psychological therapies such as these to individuals who experience CMD symptoms associated with long-term health conditions (LTC), such as diabetes. Therapists working these populations are therefore, required to follow additional competencies (National Collaborating Centre for Mental Health, 2018).

CBT and Counselling share the "common factors" of psychotherapy as described by Wampold (2015). This refers to the therapeutic relationship formed between the patient and therapist, the capacity of the intervention to provide an explanation for the targeted CMD symptoms and a mechanism of change, which can vary between approaches. The mechanisms for CBT and Counselling, as well as some of the differing practical components of delivery, are explained in turn.

CBT. Grounded in cognitive and behavioural theory, CBT conceptualises CMD symptoms through the links between an individual's thoughts, feelings, behaviours and physical sensations, which together can be summarised by formulations of various longitudinal or here-and-now formats (Fenn, 2013). Although protocol content generally corresponds to the matched CMD, CBT interventions share the aim of teaching a patient behavioural and cognitive techniques, for example, scheduling or thought-challenging. Two examples of commonly used CBT protocols derived from problem-specific formulations are the Ehlers and Clark (2000) model for PTSD and the Dugas (2004) model for Generalised Anxiety Disorder (GAD).

Regardless of protocol, CBT sessions are structured with an agenda that is largely informed by problem-orientated goals, whilst a sense of guided discovery is maintained through therapist Socratic questioning (Padesky, 1993). Depending on the CMD treated, NICE recommends between 12 and 20 sessions of CBT, although there is considerable variation between services (Clark et al., 2018; NICE, 2009; NICE, 2011).

Counselling. Although there are a multitude of Counselling approaches, the current study refers to those that belong to the humanistic tradition, which conceptualises good psychological wellbeing as arising when there is congruence between an individual's inner and outer worlds (Lambert & Erekson, 2008). Historically, Counselling of this nature has been used for both depression and anxiety disorders, however, a protocol specifically for the treatment of depression in IAPT has been developed, known as Counselling for Depression (CfD). This intervention has the aim of supporting individuals to recognise and make sense of their emotions, whilst noting the incongruence between how they feel and how they should feel based on the beliefs they hold; a concept known as self-discrepancy (Watson et al., 2010). As such, the focus of the therapist's work is to identify distressing emotional processes that arise due to self-discrepancy, such as self-criticism, and support the patient in reducing their intensity to elicit heightened mood. Meanwhile, the therapist strives to portray a relational stance, defined by authenticity, empathy and a sense of acceptance without condition (Roth et al., 2009). As a result, it is hoped that the patient is able to develop new meanings in regard to their emotions and enact positive changes in their life.

It is important to note that in some IAPT services, the remit of CfD is less stringent, whereby some therapists work integratively across modalities, with individuals experiencing either depression or anxiety disorders (Gyani et al., 2013; Pybis et al., 2017). Preliminary analyses of the current dataset have reiterated this, whereby Counselling was also offered to individuals with anxiety disorders, even though this has not been formally recommended by NICE (2011). Acknowledging this, the term ‘Counselling’ is used by the current study, encapsulating CfD and other humanistic approaches for anxiety disorders. When offered to treat depression, NICE recommends between six and ten sessions of Counselling (NICE, 2009).

Measures

Participants completed self-report measures prior to the start of every session. These measures include the PHQ-9 and GAD-7, measuring symptoms of depression and anxiety, with clinical-cut offs of ≥10 and ≥8 respectively (Kroenke et al., 2001; Spitzer et al., 2006). Above these cut-offs, patients are considered to be in ‘caseness’. Although they were not considered for the planned analyses, the IAPT Phobia Scale and the Work and Social Adjustment Scale (WSAS), measuring functional impairment, were also completed by each participant as part of the MDS (Mundt et al., 2002; National Collaborating Centre for Mental Health, 2018).

Propensity Score Matching

To control for potential confounding factors propensity score matching (PSM) was used to create matched controls for the CBT and Counselling subsamples. In line with the technique provided by Leuven and Sianesi (2003), one-to-one matching was fulfilled using the Stata package, PSMATCH2 (StataCorp, 2019). Eleven variables informed the matching process. These were age, gender, ethnicity, LTC status, psychotropic medication usage, employment status, problem descriptor (CMD diagnosis), index of multiple deprivation score (IMD), Trust of service provision, and pre-treatment scores on the PHQ-9, GAD-7 and WSAS. The continuous variables (age and the symptom measure scores) were transformed to become categorical, providing bands of age and categories of severity, ranging from low to severe, based on the measure cut-offs. A ‘missing’ category was created when this continuous data was not available. This was used as a level of these categorical variables in order to permit the use of cases during matching that would otherwise be subject to listwise deletion. To identify the best matches for each ‘treated’ observation, matching with replacement was used, whereby a participant can serve as a control or ‘best match’ on numerous occasions. The maximum difference accepted between matched participants, or ‘caliper’, was set at 0.001 to reduce bias yet ensure a sample size with sufficient power for the subsequent analyses (Lunt, 2014). As a result of PSM, 26,734 participants were removed from the final sample, most of which were participants who received CBT and were not identified as matches. Therefore, it is important to note that the final sample does not reflect the distribution of treatment types usually observed in IAPT as a result of PSM. However, there are other examples of IAPT studies that adopted a similar approach, which justified the reductions in subsample size (Saunders et al., 2020). As shown in Appendix D, all but the Trust and PHQ-9 severity variables were not significantly different between the CBT and matched Counselling group, suggesting that a good balance (indicated by Chi-Square values) of patient characteristics had been achieved via PSM.

Analysis Plan

The aims of the study were achieved with two sets of analyses using different software. The first set of analyses (GCM) were conducted on Mplus version 8, whilst the second (multinominal regressions) used IBM SPSS Statistics for Windows, Version 27.0 (Muthén & Muthén, 2017; IBM Corp, 2020). Since PSM was used, each of the below analyses were weighted using the propensity score weight variable (accounting for the same control case being used more than once). Missing data were handled as standard in Mplus, using a Full Information

Maximum-Likelihood (FIML) algorithm (Saunders et al., 2019). Meanwhile, there was no missing data as per the SPSS dataset and therefore, no further action was required.

Growth Curve Modelling. Trajectory modelling approaches vary in the number of time-points of the continuous dependent variable incorporated into the analyses. In the current study, 16 time-points of PHQ-9 and GAD-7 scores were used in an attempt to balance the risk of distorted growth curves when exceeding the mean number of sessions and the utility of the findings for encapsulating the full range of the number of sessions recommended by NICE (12 – 20).

In the first instance, latent growth curve analysis (LGCA) was used in line with Muthén & Muthén (2017) to identify the average expected response curves of PHQ-9 and GAD-7 scores over the course of each type of psychological therapy. In doing so, the individual trajectories of participants were pooled to produce an estimated mean trajectory (Curran et al., 2010). Linear and quadratic factors were fitted to the two models and treatment type (CBT and Counselling) was included as a predictor. The overall fit of the models was determined by Comparative Fit Index (CFI), Tucker-Lewis Index (TLI) and Chi-squared values.

Second, latent classes of anxiety and depression symptom trajectories were explored through application of growth mixture modelling (GMM) as described by Muthén & Muthén, (2017). Through this analysis, individual trajectories were assigned to classes on the basis of shared patterns of symptom change. To find the most suitable number of classes per measure, GMM models were fitted starting with a two-class model. The number of classes in each model was increased by one each time models were run until the Vuong-Lo-Medell-Rubin Likelihood Ratio Test (VLMR-LRT) was no longer significant ($p = \geq 0.05$) or the Bayesian Information Criterion (BIC) or Akaike Information Criterion (AIC) values exceeded that of the previous class model (Jung & Wickrama, 2008). The entropy value, for which > .80 is considered high, was also assessed for each model since higher values indicate a greater distinction between classes.

The results of the above were reported in line with the Guidelines for Reporting on Latent Trajectory Studies (GRoLTS) Checklist to ensure uniformity and high-quality, so that they can be included in systematic reviews or meta-analyses (van de Schoot et al., 2016). As part of this, the results are also depicted in a series of graphs to demonstrate the latent growth curves and trajectories of the final-class model for each measure (PHQ-9 and GAD-7). Full model statistics are provided in Appendices E and F.

Multinominal Regressions. The trajectory class that each participant was most likely to belong to was calculated as part of the GMM analyses and was extracted from Mplus into SPSS as a new variable. This variable was then used as the dependent variable of two multinominal regression models, one for each measure (PHQ-9 and GAD-7). Treatment type (CBT or Counselling) served as the predictor variable of each model fitted. The patient characteristic variables used for PSM were not included in the multinominal regressions as their influence on the class models as confounders was already controlled for through inclusion of PSM weight variable for all prior analyses.

Results

Descriptive Statistics

The final sample was comprised of 10268 participants, of which 4055 (39.5%) had CBT and 6213 (60.5%) had Counselling as their main intervention. Collectively, the average number of sessions was similar for CBT (M = 10.23, SD = 5.07) and Counselling (M = 10.07, SD = 4.49). Descriptive statistics for the sample are provided in Table 1. A more detailed summary of the patient characteristics of the CBT and Counselling subsamples, pre- and post-PSM, can be found in Appendix D. Goodness of fit indices are also provided in the appendix, demonstrating improved balance of patient characteristics between the two intervention type subsamples as a result of PSM.

Table 1

Descriptive Statistics and Frequencies for Characteristics of the Post-PSM Sample

Characteristic	Category	*n*	%
Gender	Female	7448	72.5
	Male	2520	24.5
	Intermediate	8	0.01
	Missing	292	1.8%
Ethnicity (ONS)	White	6144	59.5%
	Mixed	531	5.2%
	Asian	947	9.2%
	Black	821	8%
	Chinese	49	0.5%
	Other	542	5.3%
	Missing	1234	12%
Diagnosis	Depression	2763	26.9%
	GAD	318	3.1%
	Mixed Anxiety and Depression	365	3.6%
	OCD	4	0.0%
	PTSD	93	0.9%
	Phobic Anxiety and Panic	42	0.4%
	Unspecified Anxiety	43	0.4%
	Other	380	3.7%
	Missing	6260	61%
Characteristic	***n***	***M***	***SD***
Age	10260	43.91	15.00
Baseline PHQ-9	10206	14.73	6.24
Baseline GAD-7	10201	12.62	5.37
Baseline WSAS	9028	18.41	9.46

Latent Growth Curves

To explore whether the rate of symptom change differed between CBT and Counselling, LGCA was conducted on GAD-7 and PHQ-9 scores at the 16 time points, using the matched sample, with intervention type included as a predictor. The linear model for depression was fitted with a Chi-square of χ^2 = 5033.17 (145), p = < 0.001, alongside a CFI of 0.919 and a TLI of 0.924. Assessing the fit of the linear model for anxiety, a Chi-square value of χ^2 = 4276.94 (145), p = < 0.001 was reported, with a CFI of 0.926 and a TLI of 0.931. It was apparent that there was a quicker decrease in PHQ-9 and GAD-7 scores amongst those receiving CBT than Counselling, with β coefficients of -0.055 (95% CI [-0.084 – -0.027], < 0.001) and -0.057 (95% CI [-0.082 – -0.032], p = < 0.001) respectively. Figure 2 depicts the overall latent growth curves for PHQ-9 and GAD-7 scores, whilst the curves for the CBT and Counselling subsamples are shown in Figure 3. As shown by the curves, there was a reduction in depression and anxiety symptoms over the 16 sessions. Although linear models were fitted successfully, it was noted that these were enhanced by adding a quadratic function. Therefore, for the subsequent GMM analyses, quadratic functions were introduced for each analysis to uncover the models with the best fit possible. Appendix E provides a complete report of the fit statistics for each of the LGCA models explored (Table E1), along with model parameters (Table E2).

Figure 2

Latent Growth Curves for PHQ-9 and GAD-7 Scores over Sixteen Sessions

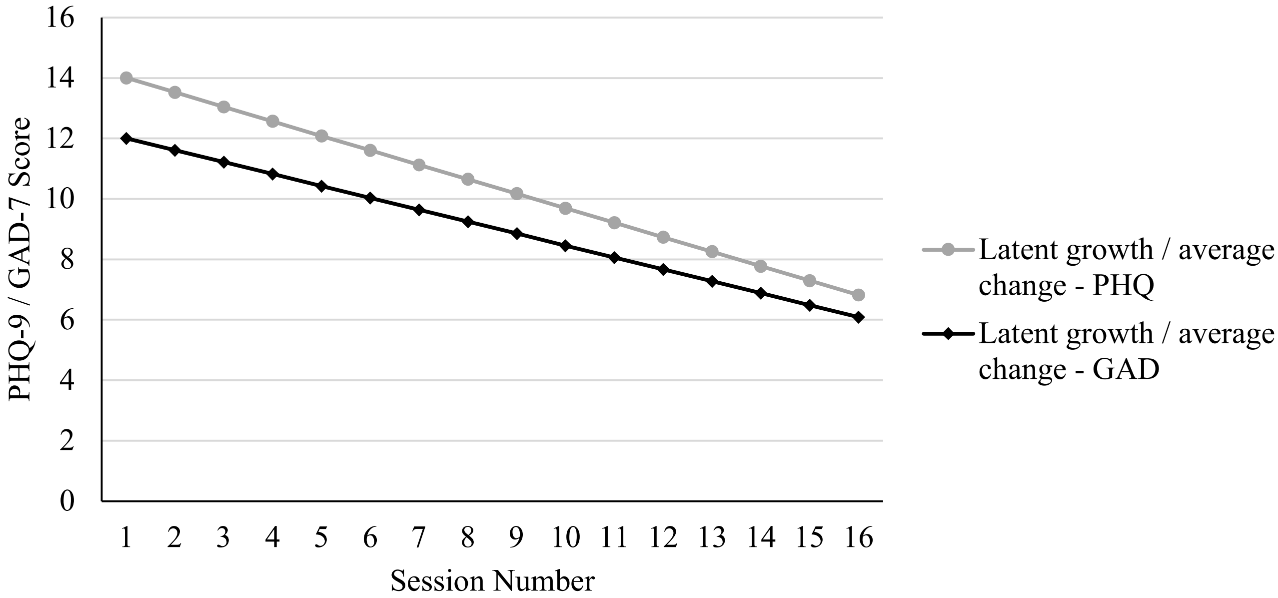
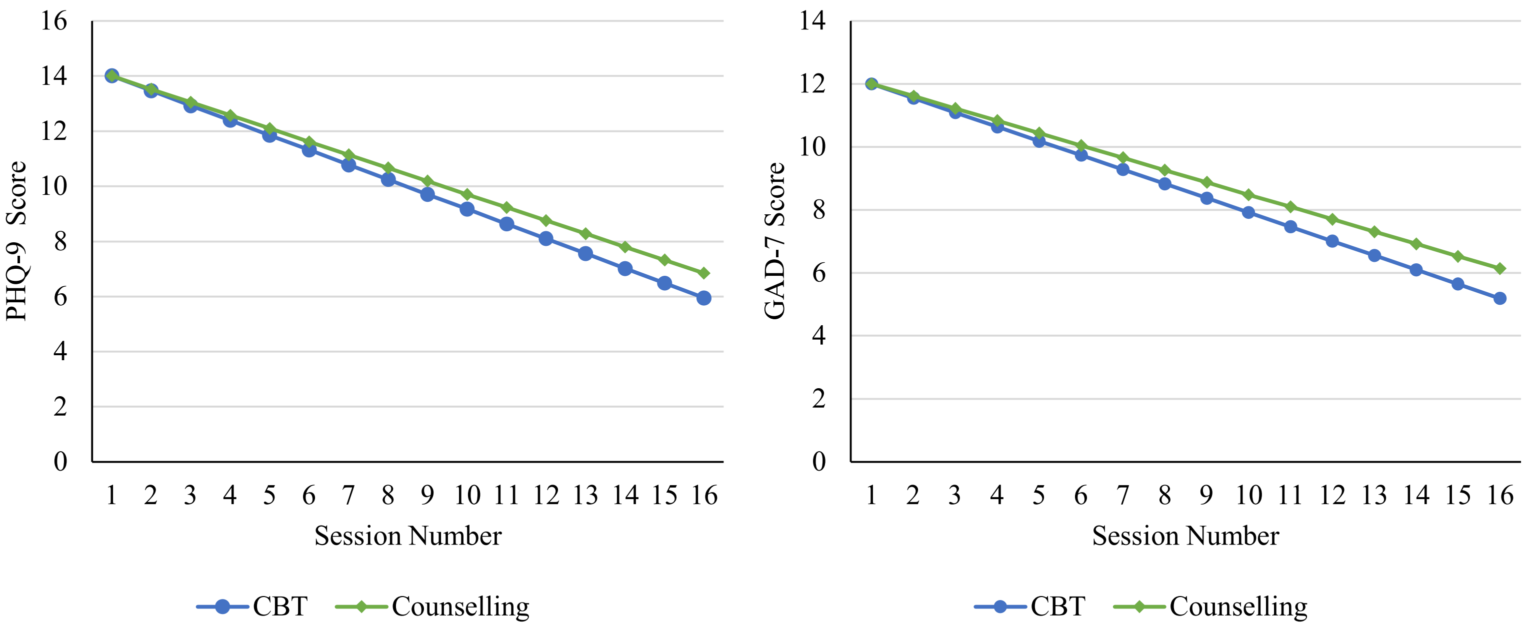


Figure 3

Latent Growth Curves for PHQ-9 and GAD-7 Scores by Intervention over Sixteen Sessions



Trajectory Classes

Using both measures, GMM was employed to explore the models of best fit for distinct trajectory classes of depression and anxiety symptom change. Since the inclusion of a quadratic factor was found to improve the fit of the latent growth curves, its use was specified during the GMM analyses. The fit indices for each class model are compared in Appendix F (Table F1). As per these comparisons, a four-class solution was chosen for both PHQ-9 and GAD-9 scores over the 16 sessions. The parameter statistics for both models are also provided in Appendix F (Table F2). As shown in Figure 4, the classes for depression and anxiety are highly similar, conducive to shared labels and descriptions. Because of this, the output of PHQ-9 classes was re-ordered to reflect that of the GAD-7 classes. These four classes are as follows:

- Class 1 – *Rapid Responders:* initial early symptom reduction, followed by a plateau and marginal increase in symptoms until the final session where overall improvement is still evident.

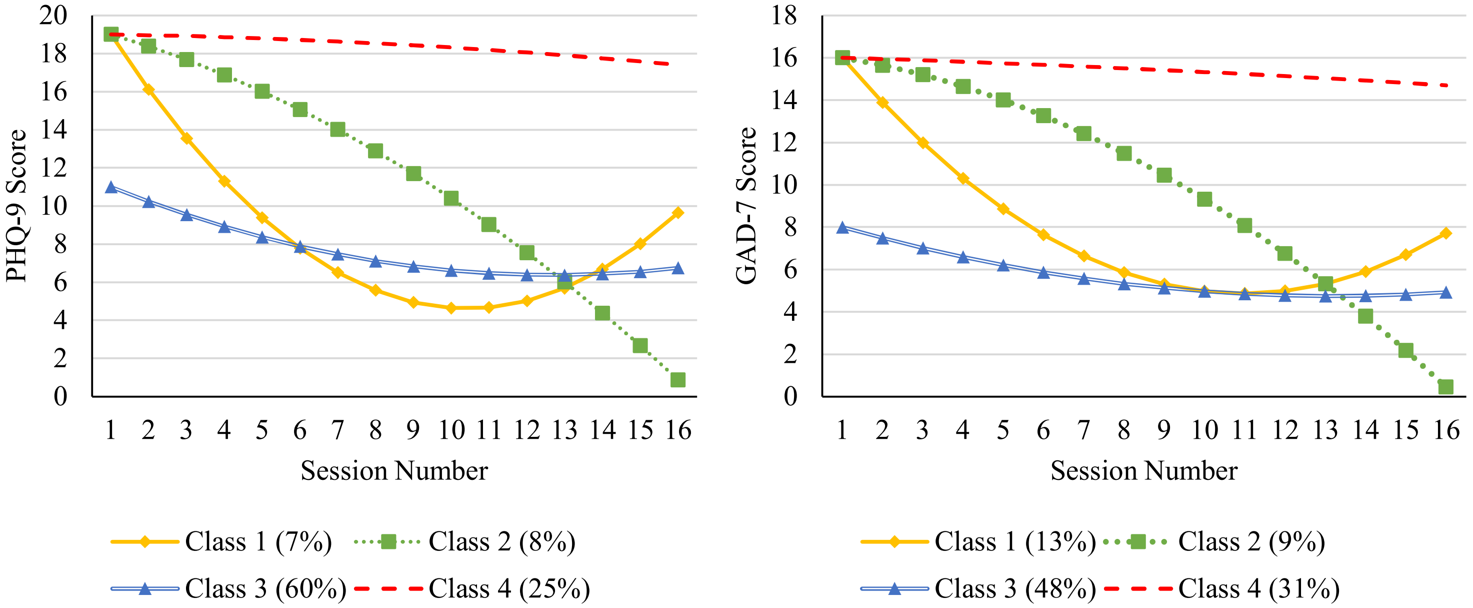
- Class 2 – *Gradual Responders*: initial moderately slow symptom reduction that hastens midway through treatment and ends up with minimal symptoms.

- Class 3 – *Low Severity Small Improvers:* symptoms of low severity pre-treatment that reduce by a small degree over the course of intervention.

- Class 4 – *Non-responders:* symptoms of high severity that do not change much over the course of treatment.

Figure 4

PHQ-9 and GAD-7 Trajectory Classes over Sixteen Sessions



These classes are best distinguished by their baseline severity, speed of symptom change and the magnitude of reduction. Concerning the former, class 3 (*Low Severity Small Improvers*) is distinct from the remaining three classes, whereby it has a lower baseline PHQ-9 and GAD-7 mean. Compared to class 3 (*Low Severity Small Improvers*) and 2 (*Gradual Responders)*, class 1 (*Rapid Responders*) has a faster rate of improvement, with a steep reduction in both measures over the first six sessions until slowing is evident from session seven. Finally, class 1 (*Rapid Responders*) and 2 (*Gradual Responders*) share a greater magnitude of symptom change, compared to the modest, yet clinically significant, improvement of class 3 (*Low Severity Small Improvers*) and the non-response of class 4 (*Non-responders*). Interestingly, deterioration was only evident for class 1 (*Rapid Responders*), which emerged between session 11 and 16 for both measures but was more marked on the PHQ-9.

Participants were most likely to belong to class 3 (*Low Severity Small Improvers*) of both the PHQ-9 (60%) and GAD-7 (48%) models, followed by class 4 (Non-responders) with membership proportions of 25% and 31%, respectively. The first class (*Rapid Responders*) of the GAD-7 model had a membership proportion of 13%, whilst the second class (*Gradual Responders*) was the smallest with 9%. Meanwhile, classes 1 (*Rapid Responders*) and 2 (*Gradual Responders*) of the PHQ-9 model had comparable membership proportions of 7% and 8%.

Co-occurrence of PHQ-9 and GAD-7 class membership was calculated (Table 2). It was apparent that the greatest degree of co-occurrence was between class 3 (*Low Severity Small Improvers*), with 43% (*n* = 5214) of the sample belonging to the class on both measures. Meanwhile, over a fifth of the sample (21.3%, *n* = 2578) belonged to class 4 (*Non-responders*) of both models. The remaining PHQ-9 and GAD-7 classes had co-occurrence rates of < 10%.

Table 2

Co-occurrences (%) of PHQ-9 and GAD-7 Trajectory Classes

PHQ-9 Class	GAD-7 Class			
	Class 1 % (*n*)	Class 2 % (*n*)	Class 3 % (*n*)	Class 4 % (*n*)
Class 1	4.6% (560)	0.5% (59)	1.5% (177)	0.7% (84)
Class 2	0.9% (108)	4.3% (519)	1.1% (133)	1.5% (182)
Class 3	6.8% (824)	3.1% (378)	43% (5214)	7% (852)
Class 4	0.6% (67)	0.7% (87)	2.4% (293)	21.3% (2578)

Associations between Intervention Type and Class Membership

Multinomial regression models were conducted with the PHQ-9 and GAD-7 classes as the dependent variables and intervention type as the independent variable. The results are presented in Table 3. Class 3 (*Low Severity Small Improvers*), the class with the largest membership proportion for both models, was used as the reference class. Considering PHQ-9 classes, Counselling was associated with a lower likelihood of belonging to class 1 (*Rapid Responders*) compared to CBT ($\beta = -.145$, $p = .043$). The same result was found for the GAD-7 classes, whereby those who received Counselling were less likely ($\beta = -.195$, $p = .001$) to be class 1 (*Rapid Responders*). In addition, Counselling was associated with a lower chance of GAD-7 class 2 (*Gradual Responders*) membership, compared to CBT ($\beta = -.152$, $p = .024$).

Table 3

Odds Ratios for PHQ-9 and GAD-9 Trajectory Class Membership by Intervention Type

Intervention	PHQ-9 Classes					
	Class 1 (n = 880)		Class 2 (n = 942)		Class 4 (n = 3025)	
	OR	(95% CI)	OR	95% CI	OR	95% CI
CBT	-	-	-	-	-	
Counselling	.865	.865 – .752	.877	.765 – 1.004	.982	.903 – 1.069
Intervention	**GAD-7 Classes**					
	Class 1 (n = 1559)		Class 2 (n = 1043)		Class 4 (n = 3696)	
	OR	(95% CI)	OR	(95% CI)	OR	(95% CI)
CBT	-	-	-	-	-	
Counselling	.822	.735 – .735	.859	.753 – .980	.953	.878 – 1.035

Note. Class 3 was used as the reference class. OR = Odds Ratio. CI = Confidence Interval.

Discussion

The current study has three main findings. First, it was evident that the overall rate of depression and anxiety symptom change differed between participants who had CBT and Counselling, with the latter group experiencing a slightly slower reduction in both types of symptoms. Second, four distinct trajectory classes were identified for depression and anxiety symptoms over the course of 16 sessions, which were highly similar in shape and form. These classes were labelled as class 1: *Rapid Responders,* class 2: *Gradual Responders*, class 3: *Low Severity Small Improvers* and class 4: *Non-responders*. The third class (*Low Severity Small Improvers*) of depression and anxiety symptoms had the greatest membership proportion, followed by class 4 (*Non-responders*). A member of either of these classes on one symptom was also likely to belong to the same class of the alternate symptom. In other words, if a participant experienced an improvement in depression symptoms from a low baseline severity (class 3*: Low Severity Small Improvers*), they may also experience the same pattern of anxiety change, or vice versa. In the same way, those who did not experience an improvement in one symptom (class 4: *Non-responders*), may also not exhibit a reduction in the other. Finally, it was found that those who received Counselling were less likely to belong to class 1 (*Rapid Responders*) of depression symptoms, as well as class 1 and 2 (*Gradual Responders)* of anxiety symptoms*,* when compared to those who had CBT.

It was apparent that depression and anxiety symptoms were slower to improve over the course of Counselling in contrast to CBT, on average. Although the difference in the rate of change of both symptoms was minor, it was observable from session three as the slopes diverged. One explanation for this may originate in the premise of “sudden gains”, whereby individuals who have CBT can experience a rapid improvement in symptoms over initial sessions (Tang & DeRubeis, 1999). It is hypothesised that these early gains are the result of a shift in cognition that prompts an upward spiral of therapeutic behavioural change. One study estimates that these sudden gains are experienced by around 41.9% of those who receive CBT for depression symptoms (Kelly et al., 2005). However, it is important to note that sudden gains have also been observed amongst individuals attending sessions of non-cognitive interventions, which has issued a debate as to whether the rapid improvements are a consequence of an adaptive cognitive style of some individuals that pre-exists intervention of any type. Irrespective of this, a meta-analysis has revealed that sudden improvements in anxiety and depression are greater amongst those who have CBT over other interventions such as Counselling (Aderka et al., 2011); a result supported by the findings of the current study

The four classes identified are consistent with the existing body of research that explored symptom trajectory classes during psychological interventions for CMDs, as reviewed in part one of this thesis. In line with the review, three classes of responder were found for both depression and anxiety, which taken together, had the greatest membership proportion over and above the fourth, non-responder class. Moreover, class 1 (*Rapid Responders*) and class 2 (*Gradual Responders*) mirrored the "Rapid Responder" and "Delayed Responder" classes as defined by the review. The finding that the classes for depression and anxiety were highly similar also echoes the review in that trajectories during treatment are often replicated regardless of the CMD symptom of interest. The current study did not find a deteriorator class during intervention, although this was not unexpected as only 4.29% of 163 classes found across 45 studies that used GCM were of this shape and form (Cole, 2021). However, marginal increases in depression and anxiety symptoms were observable from session 11 as part of class 1 (*Rapid Responders*), which is a novel finding. Only two studies of the aforementioned review reported a similar trajectory.

The first, “Remission and Recurrence”, was identified among individuals receiving high- or low-intensity CBT for depression and the second, “Rapid Response Remit”, was observed during a course of CBT for PTSD (Murphy & Smith, 2018; Wardenaar et al., 2014).

There are multiple explanations as to why an individual may appear to rapidly respond to psychological intervention as per class 1 (*Rapid Responders)*, ranging from a heightened sensitivity to the previously mentioned cognitive change integral to sudden gains or the placebo effect of therapy common factors, to a statistical artifact of regression to the mean (Nordberg et al. 2014; Tang et al., 2005; Wampold, 2015). However, the reasons for the later deterioration observed within the class are less clear. One possibility is that unlike the current study, few GCM analyses are conducted using 16 or more time-points and therefore, later deterioration for those who respond early is not identified by the models. As with an early sensitivity to the start of intervention, individuals of this class may be sensitive to the end of therapy, albeit in a less favourable manner for reasons such as failing to consolidate the initial cognitive change or fearing the termination of the therapeutic relationship (Charman & Graham, 2004). In contrast, a patient experiencing a slower rate of improvement, as characterised by class 2 (*Gradual Responders*), may not possess the individual differences responsible for sudden gains and instead, experience a ‘slow burn’ of symptom improvement. These individuals may also not experience a resurgence in symptoms nearing the end of therapy as a by-product of this slow-burn, whereby rather than confining their therapy-induced learning or skills acquisition to earlier sessions, learning is distributed over multiple sessions as per the spacing effect; something potentially more conducive to the maintenance of symptom improvements (Kim et al., 2019). Finally, class 3 (*Low Severity Small Improvers*) and 4 (*Non-responders*) trajectories are best explained by the link between baseline severity and complexity, whereby those who exhibit less severe symptoms of anxiety or depression are more likely to cross the threshold of recovery over the course of intervention, whilst those who have greater symptom severity, complexity and associated impairment are generally more resistant to the positive effect of intervention (Amati et al., 2018; Stochl et al., 2021).

Given the use of PSM to control for patient characteristics, the classes reported can be better attributed to the effect of intervention and with this, treatment type was explored as a predictor of the four classes. Building on the prior idea that the differences in the speed of symptom change between class 1 (*Rapid Responders*) and 2 (*Gradual Responders*) could be attributed to the patient's sensitivity to cognitive change and therapy-induced learning during initial sessions, it stands to reason that those who had Counselling are less likely to be members of depression class 1 (*Rapid Responders*) due to the intervention's lesser focus on formulation and skills teaching, which CBT imparts early on (Fenn & Byrne, 2013). As such, Counselling instills a slower therapeutic process, which may limit the potential for the later return of depression symptoms due to the spacing effect that may aid consolidation. However, there is currently no definitive evidence of this, and it is contradicted by the finding of Pybis and colleagues (2017) that Counselling was more effective at earlier sessions than CBT. Concerning the classes of anxiety symptoms, it is important to note that Counselling approaches are not recommended by the NICE guidelines for anxiety disorders, although as demonstrated by the current study, it is used routinely in practice (NICE, 2011). This is because Counselling approaches for anxiety have a far less expansive evidence-base, compared to the vast number of CBT RCTs. Consequently, the finding that individuals who had Counselling were less likely to belong to the responder classes with the most favourable outcomes (class 1: *Rapid Responders* and class 2: *Gradual Responders*) was not unexpected.

Limitations

The findings of the current study should be considered in recognition of a number of limitations. Firstly, although the dataset accumulated data from eight services across four London NHS Trusts, there was not a sufficient number of cases who received IPT and therefore, this intervention was excluded from the planned analyses. Very small sample sizes of additional NICE recommended interventions, such as BCT, meant that the scope of the study was limited to CBT and Counselling, despite an increase in provisions of alternative evidence-based therapies by most services over recent years (Clark, 2018). As a result, the predictive value of these other interventions in regard to the observed symptom change trajectories remains unknown. Furthermore, despite efforts to ensure that at least half of the participants' total number of sessions were either CBT or Counselling, there was no data available regarding therapist intervention fidelity. This may apply to Counselling provisions, which might reflect a range of approaches and skills in spite of a formalised offer of CfD. This was suggested by Pybis and colleagues (2017, p. 4) who stated that "many counsellors practice in an integrative manner where they bring skills and knowledge from training underpinning different forms of therapy". However, there is no direct evidence for this claim as it stands. Nonetheless, research suggests that model drift can have an adverse impact on recovery rates and in turn, could influence the trajectory classes observed over the course of sessions or their associations with the interventions themselves (Gyani et al., 2013).

The PHQ-9 and GAD-7 were the primary measures used to explore symptom trajectories during intervention. However, as of 2019/20 IAPT services have been encouraged under the Commissioning for Quality and Innovation (CQUIN) scheme to administer anxiety disorder specific measures (ADSM) during the course of intervention for specific commonly treated anxiety disorders (National Collaborating Centre for Mental Health, 2018). Although the GAD-7 has previously been reported as a “moderately good” screening tool for PTSD, panic disorder and social anxiety disorder, its use to monitor the symptom change of these disorders during intervention has been brought into question (Spitzer et al., 2006). This is because the GAD-7 does not include items specific to anxiety disorders aside from GAD and as a result, there is the potential for symptom change to be over or underestimated for some disorders, impacting patient and therapist alike (Clark, 2017). To examine whether this applies to current study, GAD-7-ADSM sensitivity analyses were considered, yet unfortunately, most ADSMs had only been used twice over the course of treatment (pre- and post-intervention), which rendered this unfeasible. Therefore, the generalisability of the trajectory classes for patients with anxiety disorders other than GAD must be interpreted with caution.

Finally, GMM was conducted using 16 time points of data, yet the average number of sessions of the sample was 10.13, with a minority (25%) of participants attending more than 12. A timeframe of 16 sessions was chosen to reflect the upper-bound of the number of recommended sessions by the NICE guidelines for the CMDs of interest and in doing so, it was hoped that the clinical utility of the findings would be optimised. Nonetheless, the interpretations made regarding the existence of the identified classes must be balanced with the possibility that the estimated trajectories approaching session 16 may be a statistical artefact due to the limited data available nearing the final data point (Muthén, 2004). Distorted trajectories for similar reasons have been reported by other studies, as well as the impact of floor effects (Lutz et al., 2005). Consequently, claims made regarding the shape of the trajectories should be perceived as solely theoretical, particularly those concerning the deterioration aspect of class 1 (*Rapid Responders*), which has not been commonly reported by other GCM studies.

Clinical and Research Implications

The classes of symptom change reported and their associations with intervention type have implications for both clinical practice and future research. In combination with the findings of other GCM studies, the results of the current study facilitate the expansion of knowledge regarding the likely responses to high-intensity psychological interventions among individuals with CMDs. As such, this newfound knowledge may assist therapists and their supervisors when making clinical decisions, for example, whether to continue with an intervention when progress is initially slow. Without the awareness of classes of change, such as Class 2 (*Gradual Responders),* patients who exhibit this pattern may have doubted their potential to improve and possibly dropped-out or may have been offered alternative therapies in the expectation that they would not otherwise recover during the course of treatment. By normalising a slower symptom trajectory, the clinician may preserve both their own hope and motivation and that of the patient, and enhance their outcomes by preventing drop-out, unnecessary termination of treatment, or an unnecessary increase in the intensity or 'aggression' of treatment e.g., with concurrent use of psychotropic medications or increased dosage of existing concurrent medications. In addition, through recognising the associations between the classes and intervention type (CBT and Counselling), the development of a more sophisticated version of the outcome feedback system devised by Delgadillo and colleagues (2018) yet sensitive to intervention type, could become possible. However, to fully grant the benefits of such an endeavor, future research is required that overcomes the previously discussed limitations of the current study. In particular, researchers should consider exploring symptom trajectories using ADSMs that may exceed the capability of the GAD-7 to capture symptom change of anxiety disorders aside GAD. Furthermore, those who have access to larger datasets may seek to explore symptom change trajectories in a sample that also includes an adequate number of cases treated with other evidence-based high-intensity psychological therapies, such as IPT, DIT, Behavioural Couples Therapy, or Collaborative Care, and consider additional treatment related factors, such as intervention fidelity, as predictors to build on the current findings.

References

Aderka, I. M., Nickerson, A., Bøe, H. S., Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: a meta-analysis. *Journal of Consulting and Clinical Psychology, 80*(1), 93–101. https://doi.org/10.1037/a0026455

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

Beck, A., Naz, S., Brooks, M., & Jankowska, M. (2019). *IAPT Black Asian and Minority Ethnic Service User Positive Practice Guide*. https://babcp.com/Portals/0/Files/About/BAME/IAPT-BAME-PPG-2019.pdf?ver=2020-06-16-004459-320

Charman, D. P., & Graham, A. C. (2004). Ending therapy: Processes and outcomes. In D. P. Charman (Ed.), *Core processes in brief psychodynamic psychotherapy: Advancing effective practice* (p. 275–288). Lawrence Erlbaum Associates Publishers

Clark D. M. (2018). Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program. *Annual review of clinical psychology, 14*, 159–183. https://doi.org/10.1146/annurev-clinpsy-050817-084833

Clark, D. M. (2017). *Why IAPT services need to use Anxiety Disorder Specific Measures (ADSM) to guide treatment and assess the overall outcome of a course of therapy*. https://www.york.ac.uk/healthsciences/pc-mis/news-articles/ADSMs_Recovery_IAPT.pdf

Clark, D. M., Canvin, L., Green, J., Layard, R., Pilling, S., & Janecka, M. (2018). Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. *The Lancet, 391*(10121), 679–686. https://doi.org/10.1016/S0140-6736(17)32133-5

Cohen, Z. D., & DeRubeis, R. J. (2018). Treatment Selection in Depression. *Annual review of clinical psychology, 14*, 209–236. https://doi.org/10.1146/annurev-clinpsy-050817-084746

Cole, C. L. (2021). *A Systematic Review of Common Mental Disorder Symptom Trajectories during Psychological Interventions and Associated Patient Characteristics* (Unpublished doctoral dissertation). University College London, UK.

Curran, P. J., Obeidat, K., & losardo, D. (2010). Twelve Frequently Asked Questions About Growth Curve Modeling. *Journal of Cognitive Development, 11*(2), 121–136. https://doi.org/10.1080/15248371003699969

Delgadillo, J., de Jong, K., Lucock, M., Lutz, W., Rubel, J., Gilbody, S., Ali, S., Aguirre, E., Appleton, M., Nevin, J., O'Hayon, H., Patel, U., Sainty, A., Spencer, P., & McMillan, D. (2018). Feedback-informed treatment versus usual psychological treatment for depression and anxiety: a multisite, open-label, cluster randomised controlled trial. *The lancet Psychiatry, 5*(7), 564–572. https://doi.org/10.1016/S2215-0366(18)30162-7

DeRubeis, R. J., Cohen, Z. D., Forand, N. R., Fournier, J. C., Gelfand, L. A., & Lorenzo-Luaces, L. (2014). The Personalized Advantage Index: translating research on prediction into individualized treatment recommendations. A demonstration. *PloS one, 9*(1). https://doi.org/10.1371/journal.pone.0083875

Dugas, M. J., Gagnon, F., Ladouceur, R., & Freeston, M. H. (1998). Generalized anxiety disorder: A preliminary test of a conceptual model. *Behavior Research and Therapy, 36*(2), 215–226. https://doi.org/10.1016/S0005-7967(97)00070-3

Ehlers, A., & Clark, C. M. (2000). A cognitive model of post-traumatic stress disorder. *Behaviour Research and Therapy, 38*(4), 319-345. https://doi.org/10.1016/s0005-7967(99)00123-0

Fenn, K., & Bryne, M. (2013). The key principles of cognitive behavioural therapy. *InnovAiT: Education and inspiration for general practice, 6*(9), 579–585. https://doi.org/10.1177/1755738012471029

Fonagy, P., Lemma, A., Target, M., O'Keeffe, S., Constantinou, M. P., Wurman, T. V., Luyten, P., Allison, E., Roth, A., Cape, J., & Pilling, S. (2020). Dynamic interpersonal therapy for moderate to severe depression: a pilot randomized controlled and feasibility trial. *Psychological Medicine, 50*(6), 1010–1019. https://doi.org/10.1017/S0033291719000928

Greater London Authority (2021). *London Datastore – Demographics*. https://data.london.gov.uk/topic/demographics

Gyani, A., Shafran, R., Layard, R., & Clark, D. M. (2013). Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy, 51*(9), 597–606. https://doi.org/10.1016/j.brat.2013.06.004

Hill, A. (2010). *The competences required to deliver effective Counselling for Depression (CfD).* https://www.ucl.ac.uk/drupal/site_pals/sites/pals/files/migratedfiles/Depression_Counselling_for_depression_clinician_s_guide.pdf

IBM Corp (2020). IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM Corp.

Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass, 2*(1), 302–317. https://doi.org/10.1111/j.1751-9004.2007.00054.x

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1–27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432

Kelly, M. A. R., Roberts, J. E., & Ciesla, J. A. (2005). Sudden gains in cognitive behavioral treatment for depression: When do they occur and do they matter? *Behaviour Research and Therapy, 43*(6), 703–714. https://doi.org/10.1016/j.brat.2004.06.002

Kim, A. S. N., Wong-Kee-You, A. M. B., Wiseheart, M., & Rosenbaum, R. S. (2019). The spacing effect stands up to big data. *Behaviour Research Methods, 51*(4), 1485–1497. https://doi.org/10.3758/s13428-018-1184-7

Kroenke, K., Spitzer, R. L., & Williams, J. B. (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., & Erekson, D. M. (2008). Positive psychology and the humanistic tradition. *Journal of Psychotherapy Integration, 18*(2), 222–232. https://doi.org/10.1037/1053-0479.18.2.222

Lambert, M. J., Whipple, J. L., Smart, D. W., Vermeersch, D., Nielsen, S. L. (2001). The Effects of Providing Therapists with Feedback on Patient Progress During Psychotherapy: Are Outcomes Enhanced?. *Psychotherapy Research, 11*(1), 49–68. https://doi.org/10.1080/713663852

Leuven, E., & Sianesi, B. (2003). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing* (Version 1.2.3). Available from: http://ideas.repec.org/c/boc/bocode/s432001

Lunt, M. (2014). Selecting an Appropriate Caliper Can Be Essential for Achieving Good Balance With Propensity Score Matching. *American Journal of Epidemiology, 179*(2), 226–235. https://doi.org/10.1093/aje/kwt212

Lutz, W., Leach, C., Barkham, M., Lucock, M., Stiles, M. B., Evans, C., Noble, R., Iveson, S. (2005). Predicting change for individual psychotherapy clients on the basis of their nearest neighbors.

Journal of Consulting and Clinical Psychology, 73(5), 904–913. https://doi.org/10.1037/0022-006X.73.5.904

Mundt, J. C., Marks, I. M., Shear, M. K., Greist, J. H. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *The British Journal of Psychiatry, 180*, 461–464. https://doi.org/10.1192/bjp.180.5.461

Murphy, D., & Smith, K. (2018). Treatment Efficacy for Veterans with Posttraumatic Stress Disorder: Latent Class Trajectories of Treatment Response and Their Predictors. *Journal of Traumatic Stress, 31*(5), 753–763. https://doi.org/10.1002/jts.22333

Muthén L. K., & Muthén, B. O. (2017). *Mplus User's Guide. Eighth Edition.* https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf

Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of Quantitative Methodology for the Social Sciences* (pp. 345–368). Newbury Park, CA: Sage Publications

National Collaborating Centre for Mental Health (2018). *The Improving Access to Psychological Therapies (IAPT) Pathway for People with Long-term Physical Health Conditions and Medically Unexplained Symptoms.* https://www.england.nhs.uk/wp-content/uploads/2018/03/improving-access-to-psychological-therapies-long-term-conditions-pathway.pdf

National Collaborating Centre for Mental Health (2020). *The Improving Access to Psychological Therapies Manual.* https://www.england.nhs.uk/wp-content/uploads/2020/05/iapt-manual-v4.pdf

NHS Digital (2020). *Psychological Therapies, Annual report on the use of IAPT services 2019–20.* https://files.digital.nhs.uk/B8/F973E1/psych-ther-2019-20-ann-rep.pdf

NICE (2009). *Depression in adults: recognition and management Clinical guideline [CG90].* https://www.nice.org.uk/guidance/cg90

NICE (2011). *Common mental health problems: identification and pathways to care Clinical guideline [CG123]*. https://www.nice.org.uk/guidance/cg123

NICE (2011). *Generalised anxiety disorder and panic disorder in adults: management Clinical guideline [CG113]*. https://www.nice.org.uk/guidance/cg113

Nordberg, S. S., Castonguay, L. G., Fisher, A. J., Boswell, J. F., & Kraus, D. (2014). Validating the Rapid Responder Construct Within a Practice Research Network. *Journal of Clinical Psychology, 70*(9), 1–18. https://doi.org/10.1002/jclp.22077

Office for National Statistics (2012). *2011 Census: Key Statistics for England and Wales, March 2011.* https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/2011censuskeystatisticsforenglandandwales/2012-12-11

Padesky, C. A. (1993). *Socratic Questioning: Changing Minds or Guiding Discovery?.* Invited keynote address presented at the 1993 European Congress of Behaviour and Cognitive Therapies, London. https://padesky.com/newpad/wp-content/uploads/2012/11/socquest.pdf

Power, M. J. & Freeman, C. (2012). A Randomized Controlled Trial of IPT Versus CBT in Primary Care: With Some Cautionary Notes About Handling Missing Values in Clinical Trials. *Clinical Psychology and Psychotherapy, 19*(2), 159–169. https://doi.org/10.1002/cpp.1781

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

Ram, N., & Grimm, K. J. (2009). Growth Mixture Modeling: A Method for Identifying Differences in Longitudinal Change Among Unobserved Groups. *International Journal of Behavioural Development, 33*(6), 565–576. https://doi.org/10.1177/0165025409343765

Roth, A., & Pilling, S. (2007). *The competences required to deliver effective cognitive and behavioural therapy for people with depression and with anxiety disorders.* https://www.ucl.ac.uk/drupal/site_pals/sites/pals/files/migrated-files/Backround_CBT_document_-_Clinicians_version.pdf

Roth, A., Fonagy, P. (2005). *What works for whom: A critical review of psychotherapy research* (2nd ed.). Guilford Publications.

Roth, A.D., Hill, A., & Pilling, S. (2009). *The competences required to deliver effective humanistic psychological therapies.* https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.572.6986&rep=rep1&type=pdf

Saunders, R., Buckman, J. E. J., Cape, J., Fearon, P., Leibowitz, J., & Pilling, S. (2019). Trajectories of depression and anxiety symptom change during psychological therapy. *Journal of Affective Disorders, 249*, 327–335. https://doi.org/10.1016/j.jad.2019.02.043

Saunders, R., Buckman, J., & Pilling, S. (2020). Latent variable mixture modelling and individual treatment prediction. *Behaviour research and therapy, 124*(103505). https://doi.org/10.1016/j.brat.2019.103505

Saunders, R., Cape, J., Leibowitz, J., Aguirre, E., Jena, R., Cirkovic, M., Wheatley, J., Main, N., Pilling, S., & Buckman, J. (2020). Improvement in IAPT outcomes over time: are they driven by changes in clinical practice?. *Cognitive behaviour therapist, 13*. https://doi.org/10.1017/S1754470X20000173

Schulz, K. F., Altman, D. G., Moher, D., & CONSORT Group (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC medicine, 8*(18). https://doi.org/10.1186/1741-7015-8-18

Spitzer, R. L., Kroenke, K., Williams, J. B. W, Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine, 166*(10), 1092–2097. https://doi.org/10.1001/archinte.166.10.1092.

StataCorp (2019). Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC.

Steel, Z., Marnane, C., Iranpour, Chey, T., Jackson, J. W., Patel, V., & Silove, D. (2014). The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *International Journal of Epidemiology, 43*(2), 476–493. https://doi.org/10.1093/ije/dyu038

Stochl, J., Soneson, E., Stuart, F., Fritz, J., Walsh, A. E. L., Croudance, T., Hodgekins, J., Patel, U., Russo, D. A., Knight, C., Jones, P. B., & Perez, J. (2021). Determinants of patient-reported outcome trajectories and symptomatic recovery in Improving Access to Psychological Therapies (IAPT) services. *Psychological Medicine, 8*, 1–10. https://doi.org/10.1017/S0033291720005395

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology, 67*(6), 894–904. https://doi.org/10.1037/0022-006X.67.6.894

Tang, T. Z., DeRubeis, R. J., Beberman, R., & Pham, T. (2005). Cognitive Changes, Critical Sessions, and Sudden Gains in Cognitive-Behavioral Therapy for Depression. *Journal of Consulting and Clinical Psychology, 73*(1), 168–172. https://doi.org/10.1037/0022-006X.73.1.168

van de Schoot, R., Sijbrandij, M., Winter, S. D., Depaoli, S., & Vermunt, J. K. (2016). The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. *Structural Equation Modeling: A Multidisciplinary Journal, 24*(3), 451–467. https://doi.org/10.1080/10705511.2016.1247646

Wakefield, S., Kellett, S., Simmonds-Buckley, M., Stockton, D., Bradbury, A., & Delgadillo, J. (2020). Improving Access to Psychological Therapies (IAPT) in the United Kingdom: A systematic review and meta-analysis of 10-years of practice-based evidence. *British Journal of Clinical Psychology, 60*(1), 1-37. https://doi.org/10.1111/bjc.12259

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110*(20), 639–644. https://doi.org/10.2466/02.07.17.PR0.110.2.639-644

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry, 14*(3), 270–277. https://doi.org/10.1002/wps.20238

Wardenaar, K. J., Conradi, H. J., & De Jonge, P. (2014). Data-driven course trajectories in primary care patients with major depressive disorder. *Depression and Anxiety, 31*(9), 778–786. https://doi.org/10.1002/da.22228

Watson, N., Bryan, B. C., & Thrash, T. M. (2010). Self-discrepancy: Comparisons of the psychometric properties of three instruments. *Psychological Assessment, 22*(4), 878–892. https://doi.org/10.1037/a0020644

Part 3: Critical Appraisal

A Critical Appraisal of Part One (Literature Review) and Two (Empirical Paper) of the Doctoral Thesis

Introduction

The following reflections are stated as part of a critical appraisal of the systematic review and empirical project, which comprise the overall doctoral thesis that explored differential trajectories, or ‘classes’, of common mental disorder (CMD) symptoms during psychological interventions. The review sought to investigate whether the trajectory classes described by the existing research literature differed on the basis of the psychological intervention offered, the CMD treated and the medium of intervention delivery, as well as summarising the patient characteristics that were associated with these classes. Meanwhile, the empirical project addressed gaps in the literature by exploring whether the observable classes of CMD symptom change were associated with the intervention types of Cognitive Behavioural Therapy (CBT) and Counselling.

Concerning the systematic review process, the following are discussed in turn. The necessity of outcome feedback technologies in face of flawed clinical intuition, the Guidelines for Reporting on Latent Trajectory Studies (GRoLTS; van de Schoot et al., 2017) and the relevance of the “Dodo Bird Verdict” to trajectory classes (Rosenzweig, 1936). A commentary is then provided regarding the empirical project, covering the topics of NICE guidelines non-compliance, the process of learning an advanced statistical approach and the experience of working with a large dataset. Finally, an account is given of undertaking a doctoral thesis during the Covid-19 pandemic and the pressures this entailed.

Reflections on the Systematic Review

Routine Outcome Monitoring (ROM)

A majority of the studies included in the review cited the use of ROM during psychological intervention and discussed the scope of growth curve modelling (GCM) studies to enhance its clinical utility. Despite this apparent rise in popularity among researchers over the past decade, few services I have worked for on placement during clinical training have made use of ROM technologies. Meanwhile, in IAPT services where ROM is used, uptake amongst clinicians appeared to be inconsistent. Having familiarised myself with the advantages it entails through reviewing the relevant literature, I have questioned whether a "scientist-practitioner gap" is emerging that renders ROM a 'missed opportunity' (Cautin, 2011). To explore this, I have considered some of the reasons why clinicians may not use it when available.

Firstly, I recognise from my own clinical practice that ROM requires significant practical investments in terms of time, additional training and administration, which may feel unsustainable due to high workloads that confer a risk of burnout (Westwood et al., 2017). Secondly, clinicians may hold attitudes or beliefs about the use of ROM, which lessen their motivation to implement it. One example of this I have encountered, has come from a colleague who suggests that ROM seeks to "objectify the therapeutic experience, which is otherwise inherently subjective". In extension of this, some clinicians believe that focusing on the 'numbers' may interfere with the therapeutic relationship (Youn et al., 2012). Likewise, it would be expected that some clinicians would hold concerns about how the data could be used, for example, the application of ROM for measuring staff performance. Finally, hesitancy in using ROM may stem from the philosophical stance that the clinical judgment of therapists is superior to that of the algorithms adopted by ROM technologies or that 'free will' is eroded by its probabilistic nature (Hatfield & Ogles, 2004). However, as stated throughout the thesis, investigations into the accuracy of clinical intuition have shown it to be biased to the extent that ROM provides a useful tool to counteract it (Walfish et al., 2012).

Taken together, there are substantial reasons as to why a scientist-practitioner gap may be transpiring in regard to the development and use of ROM, some of which I can relate to. Given the advantages it entails, I am motivated to encourage and support others to use ROM in their clinical practice as I progress throughout my career. However, bottom-up efforts will not be sufficient in narrowing the gap. Instead, services must provide adequate training that highlights the benefits of using ROM for clinicians and patients alike, whilst organisational change is required to ensure a workplace culture that welcomes its implementation.

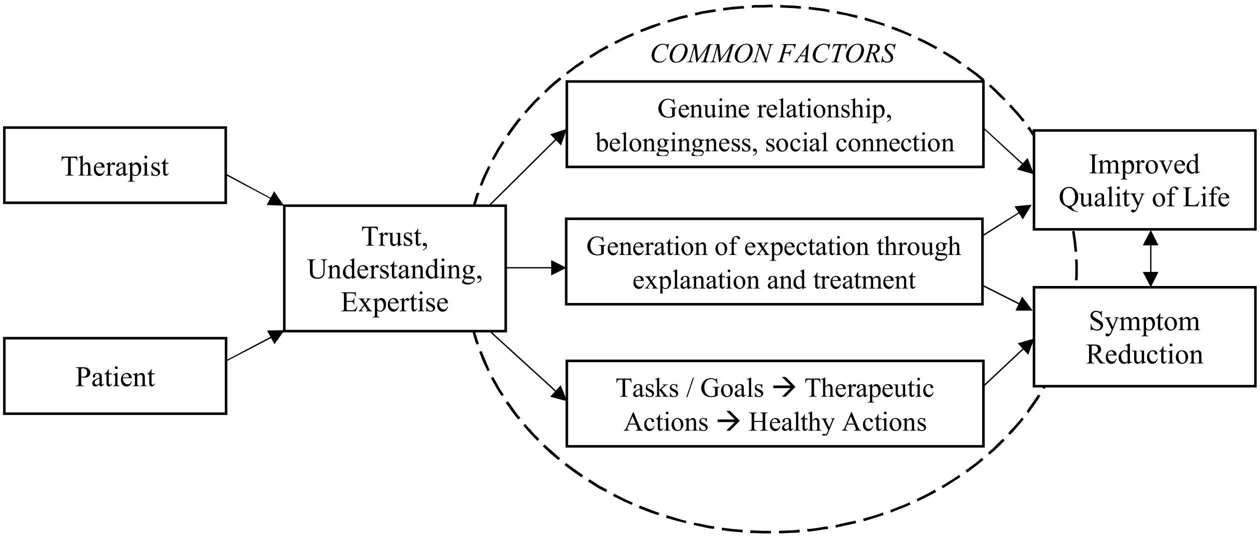
The "Dodo Bird Verdict"

Prior to clinical training I held the reasonably strong belief that CBT was superior to other forms of psychological intervention, perhaps due to my background of working as a low-intensity CBT practitioner in IAPT. When reviewing studies of varying interventions as part of the systematic review, I found that many of my assumptions about the likely findings were disputed. In particular, I had expected to uncover variations in symptom class trajectories between the interventions of interest, with a greater proportion of favourable outcomes evident amongst those receiving CBT. However, this was not the case and my eagerness to uncover substantially different mechanisms of change, made apparent by distinct observable trajectories, was challenged (Kazdin, 2007). Furthermore, these classes appeared to be associated with patient characteristics, such problem severity and complexity, whilst factors related to the intervention itself seemed to serve a minimal role.

When interpreting this finding, I was reminded of the "Dodo Bird Verdict", a reference to the Dodo of 'Alice in Wonderland' fame who declared, "Everybody has won, and all must have prizes.", when asked who had won a race. This notion of equivalence was applied to the effectiveness of psychological interventions by Rosenzweig (1936), suggesting that the technical differences of interventions have a minimal impact on overall patient outcomes. Instead, the fundamental ingredients of psychological therapy aside from the intervention itself were emphasised. These patient-therapist variables became known as the "common factors", defined by the contextual model (Figure 1; Wampold & Imel, 2015) as: the genuine relationship formed between therapist and patient, the patient expectation of improved mental health through engagement with the intervention and a specific treatment component that enacts adaptive change.

Figure 1

A Reproduced Graphical Representation of the Contextual Model of Common Factors (Wampold & Imel, 2015)



Upon consideration of this model in regard to the findings of the systematic review, I have come to acknowledge that my initial view of CBT as outright 'superior' reflects a position of naivety; something that was thankfully transformed over the course of doctoral training. In contrast to my former stance, I find myself accepting the complexity of psychological interventions and have warmed to the idea of common factors having a potentially greater role in determining patient outcomes than the underpinning mechanisms of differing interventions.

Quality Assessment with the GRoLTS Checklist

The GRoLTS checklist was employed to assess the reporting standards of the trajectories identified across the 45 studies included in the review. This process was highly time consuming, since 16 items had to be applied to each individual study and a considerable amount of time was spent trying to comprehend the terminology used to describe the concepts assessed by the checklist. Based on the scores of each item, the following observations were made about the general quality of trajectory reporting, indicative of areas of improvement for researchers who may wish to use the statistical methods applicable to the review. Most importantly, none of the studies provided graphical representations of the estimated mean trajectories for each of the class models explored (item 14b and c). Despite having a different subject matter, another review of 200 trajectory studies found that the item was also overlooked in every instance (Erosheva et al., 2014). Item three was also commonly disregarded by the studies, whereby the mechanism of missing data was not reported, with most taking the software default setting for granted. Finally, very few studies made their syntaxes freely available.

Despite implying that the majority of studies included in the review had room for improvement when reporting trajectories, the use of the GRoLTS checklist was helpful for my own development. Firstly, by acknowledging the common pitfalls of studies in the area of trajectory research, I was able to anticipate the ways in which I could enhance the quality of my own empirical project by paying particular attention to the checklist items commonly overlooked when writing up the results. Likewise, as a result of having to gain a new understanding of trajectory modelling concepts when applying its items, the GRoLTS checklist facilitated my ability to interpret the outputs of my analyses, which saved me both time and effort.

Reflections on the Empirical Project

NICE Guideline Non-compliance

Having conducted the empirical project using data from IAPT services where intervention provisions are intended to be informed by NICE guidance, it has been eye-opening to present data that shows this is not always the case. Because of this, I have found myself pondering the reasons why services and therapists may occasionally overlook the guidance. These are discussed in turn, alongside two examples of non-compliance made evident by the empirical project.

Firstly, when defining the sample, cases were excluded if they did not receive either CBT or Counselling for at least half of their total number of sessions. I was surprised to learn that a considerable proportion of participants received sessions of either of these interventions amidst another primary treatment, ranging from Interpersonal Therapy (IPT) to Eye Movement Desensitisation Reprocessing (EMDR) therapy. Therefore, the therapists who offered CBT or Counselling to these individuals cannot be viewed as upholding 'protocol fidelity'; an expectation of the IAPT programme and NICE (Clark, 2011). A common criticism of protocols is that some clinicians consider them to be "cookbooks", whereby their apparent rigidity can hinder the therapist's sensitivity to a patient's individual needs (Kendall & Frank, 2018, p. 8). Because of this, some of the therapists who delivered interventions to participants in the sample, may have chosen to use an integrative approach that included sessions of differing modalities, rather than offering a 'purist' intervention. Some also argue that protocols are a barrier to an effective therapeutic alliance due to the pressures experienced by the therapist to adhere to a particular model through ‘scripted’ use of language. However, there is limited research to evidence this claim (Kendall & Frank, 2018).

There is also the commonly held belief that protocols put the patient’s difficulties in a ‘box’ and do not sufficiently address co-occurring problems. Therefore, some clinicians may choose to draw on elements of differing interventions, depending on the problem the patient brings to the session, rather than keeping to protocol. However, reviews of recent randomised controlled trials (RCT) that have informed the development of protocols have not supported this view, highlighting that their inclusion criteria have generally become less stringent, allowing for a degree of comorbidity usually observed in routine practice (Schinder et al., 2011). When compared, there is also little difference in the outcomes of those who do and do not meet the more rigorous inclusion criteria of typical RCTs (Post et al., 2013). Therefore, whilst it is true that protocols generally target a main problem, it stands to reason that protocols informed by RCTs can accommodate the needs of those with and without comorbidity.

Perhaps the most significant example of NICE non-compliance uncovered by the empirical project was the provision of Counselling for individuals with anxiety disorders and PTSD. As it stands NICE does not recommend Counselling for these populations. However, as was the case for the study sample, there are multiple examples of IAPT services delivering the intervention despite this (Gyani et al., 2013; Pybis et al., 2017). Having worked in multiple IAPT services, I have also observed this happen first-hand, with little to no acknowledgment of the NICE guidelines. Although this applies to a minority of patients, it is concerning to discover that the recovery rates of those who received Counselling for an anxiety disorder were considerably lower (39.7%) than those who had CBT (54.2%) as their primary intervention (Gyani et al., 2013). To understand the reasoning behind this, I consulted a colleague who offers Counselling in this manner. They explained that they believed the NICE guidelines are lagging behind the "innovative" care provided by services, with Counselling provisions for anxiety being an example of this. Furthermore, they raised their concern that NICE overlooks extensive research in favour of Counselling for anxiety, as well as evidence that suggests the equivalence of psychological interventions more broadly due to common factors; a debate covered throughout this thesis. Since Counselling is less protocolised in nature compared to CBT and therefore, less 'testable', it is also understandable why fewer RCTs have been conducted.

Reflecting on the two examples of NICE non-compliance uncovered by my empirical project has been enlightening and has highlighted some of the reasons why clinicians may not choose to follow protocols, as well as some of the prevalent criticisms of NICE (Mollon, 2009). This has been helpful to clarify my own positioning, which is ultimately in favour of NICE and protocolised interventions. Above all, I welcome the premise that the application of protocols and the NICE guidelines is best performed when there is "flexibility within fidelity", achieving a balance between patient needs and the evidence-base (Kendall & Frank, 2018, p. 1).

Learning to use Growth Curve Modelling (GCM) Analyses

The empirical project required me to learn and apply the two GCM techniques of latent growth curve analysis (LGCA) and growth mixture modelling (GMM); both of which can be considered 'advanced' statistical approaches (van der Nest et al., 2020). Because these were not covered by the course teaching, their use, along with the statistical concepts that underlie them, were largely self-taught with guidance from my supervisor. On reflection, this process was the greatest challenge I faced during the completion of the empirical project and required a 'trial-and-error' attitude when troubleshooting the issues that arose. This process provides an example of 'self-directed learning', occurring when "individuals take the initiative, with or without the help of others, in diagnosing their learning needs, formulating learning goals, identifying resources for learning, choosing and implementing appropriate learning strategies, and evaluating learning outcomes" (Knowles, 1975, p. 18). As part of this, I found watching YouTube tutorials on MPlus particularly beneficial, as well as searching for solutions to syntax errors on online forums.

My supervisors' approach was also conducive to adult learning by adopting a 'coaching' style, characterised by the transfer of experience and skills through balancing supervisee self-directed learning with informal training, whilst providing constructive feedback on tasks that were collaboratively planned (Dunst & Trivette, 2012). During this process, a trainee in my cohort was also having to learn the same GCM analyses. Therefore, it was incredibly helpful to discuss the process with them, which permitted peer learning and support; something I have really valued over the course of the empirical project (Boud et al., 2001). Despite the challenges I faced when learning and implementing the GCM approaches, I have pushed my boundaries and feel proud for doing so. The experience has also highlighted to me the utility of the approach for psychological research, and I will endeavour to use it in the future when possible.

Large Datasets and Feelings of Dehumanisation

GCM analyses, such as growth mixture modelling (GMM), require large sample sizes to acquire valid findings (Kim, 2014). Prior to applying the exclusion criteria, the dataset used for the empirical project consisted of approximately half a million participants and even after the removal of a large number of cases, which couldn't be included in the analyses, the dataset became the largest I have handled to date. Amidst moments when my PC crashed due to its size, I realised how many people's lives were represented in the rows and columns of the dataset; something that could easily be forgotten when fixated on the processes involved in research. This experience was repeated when reading the statistical outputs, which were fairly short given the thousands of people who provided data to inform them; each with their own narratives of accessing psychological intervention, which went unheard by the quantitative methods I was employing. In many ways, this experience of overlooking the person behind the data by favouring their objective, rather than subjective, experience can be likened to dehumanisation; a process defined as “the denial of qualities associated with meaning, interest, and compassion” (Barnard, 2001, p. 98). One of the cognitive mechanisms proposed as having a role in dehumanisation among individuals handling datasets is a hindered ability to ascribe an identify to members of a population due to a sense of psychological distance, which can grow as sample size increases (Haslam, 2006).

It is therefore, somewhat of an irony that research such as my empirical project, which seeks to pave the way to more personalised interventions, can feel so impersonal through utilising large datasets. As a result, I experienced occasional feelings of cognitive dissonance when my values of person-centredness as a psychologist conflicted with a statistical approach that felt so detached from patient narratives, especially since I did not collect the data myself (Brehm et al., 1962). Although I view this as something that is inherent to the chosen analytic method and not necessarily problematic when the findings are used to better lives, I found myself thinking of ways to limit the potential sense dehumanisation when working with big datasets. In particular, I considered the possibility of incorporating qualitative research approaches to complement the use of GMM, by interviewing individuals belonging to each of the trajectory classes identified. As a result of a mixed-methods approach such as this, the results could be considered from the patient perspective and a richer understanding of what it is like to be either a ‘responder’ or ‘non-responder’ to psychological intervention could be ascertained (Regnault et al., 2018). Moreover, this may 'bring life' to the classes, combating the sense of detachment the researcher may have from the participants constituting the large dataset.

Undertaking a Doctoral Thesis during the Covid-19 Pandemic

The majority of the thesis was completed during the Covid-19 pandemic, which presented me with substantial pressures, both personal and professional, besides the academic demands I was already facing as a trainee clinical psychologist. Undergoing this shared experience of adjusting to the "new normal" felt heavy and distressing, yet I also became aware of and thankful for the privileges I am afforded in life. It also became clear that the pandemic was a matter of living through "the same storm, but in different boats", with frontline NHS staff facing unprecedented stressors in particular. Acknowledging this, I embarked on the process of establishing psychological support provisions for hospital staff, drawing on relevant prior experiences (Cole et al., 2020). Although this project felt like it could have warranted a thesis in its own right, I found myself experiencing an unexpected phase of personal growth; something that assisted the thesis process by giving me a "second wind" when my motivation was admittedly low at the midpoint of clinical training. It also demonstrated to me the dynamic role clinical psychologists can play in responding to crises such as Covid, giving me further incentive to reach the finish line and qualify through submission of the thesis.

Conclusion

The reflections I have put forward in this critical appraisal concern the experience of conducting a systematic review and empirical project in fulfilment of my Doctorate in Clinical Psychology. These have touched upon key debates in clinical practice, including the use of ROM, the "Dodo Bird Verdict" and the recommendations of the NICE guidelines. They have also highlighted key learning points regarding the use of a quality assessment tool for trajectory research (the GRoLTS checklist) and the application of GCM analyses on a large dataset. Finally, these reflections were contextualised through provision of an account of completing the thesis during the Covid-19 pandemic.

References

Barnard, A. (2001). On the Relationship between Technique and Dehumanization. In Locsin, R. (Eds.), *Advancing Technology, Caring, and Nursing* (pp. 96-105). Westport, CT: Auburn House. http://hdl.handle.net/10822/546880

Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance*. John Wiley & Sons Inc. https://doi.org/10.1037/11622-000

Cautin R. L. (2011). Invoking history to teach about the scientist-practitioner gap. *History of psychology, 14*(2), 197–203. https://doi.org/10.1037/a0022897

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

Cole, C. L., Waterman, S., Stott, J., Saunders, R., Buckman, J., Pilling, S., & Wheatley, J. (2020). Adapting IAPT services to support frontline NHS staff during the Covid-19 pandemic: the Homerton Covid Psychological Support (HCPS) pathway. *Cognitive behaviour therapist, 13*. https://doi.org/10.1017/S1754470X20000148

Dunst, C. J., & Trivette, C. M (2012). Moderators of the Effectiveness of Adult Learning Method Practices. *Journal of Social Sciences, 8*(2), 143–148. https://doi.org/10.3844/jssp.2012.143.148

Gyani, A., Shafran, R., Layard, R., & Clark, D. M. (2013). Enhancing recovery rates: lessons from year one of IAPT. *Behaviour research and therapy, 51*(9), 597–606. https://doi.org/10.1016/j.brat.2013.06.004

Haslam N. (2006). Dehumanization: an integrative review. *Personality and social psychology review: an official journal of the Society for Personality and Social Psychology*, 10(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4

Hatfield, D. R., & Ogles, B. M. (2004). The Use of Outcome Measures by Psychologists in Clinical Practice. *Professional Psychology: Research and Practice, 35*(5), 485–491. https://doi.org/10.1037/0735-7028.35.5.485

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1–27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432

Kendall, P. C., & Frank, H. E. (2018). Implementing evidence-based treatment protocols: Flexibility within fidelity. *Clinical psychology: a publication of the Division of Clinical Psychology of the American Psychological Association*, 25(4). https://doi.org/10.1111/cpsp.12271

Knowles, M. S. (1975). Self-directed learning. Chicago: Follett Publishing.

Boud, D., Cohen, R., & Sampson, J. (2001). *Peer learning in higher education: Learning from and with each other*. Kogan Page Limited.

Mollon, P. (2009). The NICE guidelines are misleading, unscientific, and potentially impede good psychological care and help. *Psychodynamic Practice, 15*(1), 9–24. https://doi.org/10.1080/14753630802614457

Post, P. N., de Beer, H., & Guyatt, G. H. (2013). How to generalize efficacy results of randomized trials: recommendations based on a systematic review of possible approaches. *Journal of evaluation in clinical practice, 19*(4), 638–643. https://doi.org/10.1111/j.1365-2753.2012.01888.x

Pybis, J., Saxon, D., Hill, A., & Barkham, M. (2017). The comparative effectiveness and efficiency of cognitive behaviour therapy and generic counselling in the treatment of depression: evidence from the 2nd UK National Audit of psychological therapies. *BMC psychiatry, 17*(1), 215. https://doi.org/10.1186/s12888-017-1370-7

Regnault, A., Willgoss, T., & Barbic, S. (2018). Towards the use of mixed methods inquiry as best practice in health outcomes research. *Journal of Patient-Reported Outcomes, 2*(19). https://doi.org/10.1186/s41687-018-0043-8

Rosenzweig, S. (1936). Some Implicit Common Factors in Diverse Methods of Psychotherapy. *American Journal of Orthopsychiatry, 6*, 412–415. https://doi.org/10.1111/j.1939-0025.1936.tb05248.x

Schindler, A. C., Hiller, W., & Witthöft, M. (2011). Benchmarking of cognitive-behavioral therapy for depression in efficacy and effectiveness studies--how do exclusion criteria affect treatment outcome?. *Psychotherapy research: journal of the Society for Psychotherapy Research, 21*(6), 644–657. https://doi.org/10.1080/10503307.2011.602750

van de Schoot, R., Sijbrandij, M., Winter, S. D., Depaoli, S., & Vermunt, J. K. (2017). The GRoLTS-checklist: Guidelines for reporting on latent trajectory studies. *Structural Equation Modeling, 24*(3), 451–467. https://doi.org/10.1080/10705511.2016.1247646

van der Nest, G., Passos, V. L., Candel, M. J. J. M., & van Breukelen, G. J. P. (2020). An overview of mixture modelling for latent evolutions in longitudinal data: Modelling approaches, fit statistics and software. *Advances in Life Course Research, 43*. https://doi.org/10.1016/j.alcr.2019.100323

Walfish, S., McAlister, B., O'Donnell, P., & Lambert, M. (2012). An investigation of self-assessment bias in mental health providers. *Psychological Reports, 110*(20), 639–644. https://doi.org/10.2466/02.07.17.PR0.110.2.639-644

Wampold B. E. (2015). How important are the common factors in psychotherapy? An update. *World psychiatry: official journal of the World Psychiatric Association (WPA), 14*(3), 270–277. https://doi.org/10.1002/wps.20238

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge/Taylor & Francis Group.

Westwood, S., Morison, L., Allt, J., & Holmes, N. (2017). Predictors of emotional exhaustion, disengagement and burnout among improving access to psychological therapies (IAPT) practitioners. *Journal of mental health, 26*(2), 172–179. https://doi.org/10.1080/09638237.2016.1276540

Youn, S. J., Kraus, D. R., & Castonguay, L. G. (2012). The treatment outcome package: facilitating practice and clinically relevant research. *Psychotherapy, 49*(2), 115–122. https://doi.org/10.1037/a0027932

Appendix A

Table A1

A Summary of Psychological Interventions for CMDs as Recommended by the NICE Guidelines (Clark et al., 2018)

CMD	Step of Care Provision	Recommended Intervention
Depression	Step 2: Low-intensity psychological intervention	Individual Guided Self-Help based on CBT, Computerised CBT, Behavioural Activation, Structured Group Physical Activity Program
	Step 3: High-intensity psychological intervention	CBT or IPT (with or without medication), Behavioural Activation (BA), Couples Therapy, Counselling for Depression, Brief psychodynamic therapy *For relapse prevention:* CBT, Mindfulness based Cognitive Therapy
Panic Disorder	Step 2: Low-intensity psychological intervention	Guided Self-Help based on CBT, Psychoeducational Groups, Computerised CBT
	Step 3: High-intensity psychological intervention	CBT
Generalised Anxiety Disorder (GAD)	Step 2: Low-intensity psychological intervention	CBT
	Step 3: High-intensity psychological intervention	Guided Self-Help based on CBT, Psycho-educational Groups, Computerised CBT
Social Anxiety Disorder (SAD)	Step 2: Low-intensity psychological intervention	None

	Step 3: High-intensity psychological intervention	CBT
Post-traumatic Stress Disorder (PTSD)	Step 2: Low-intensity psychological intervention	None
	Step 3: High-intensity psychological intervention	CBT, EMDR
Obsessive Compulsive Disorder (OCD)	Step 2: Low-intensity psychological intervention	Guided Self-Help based on CBT
	Step 3: High-intensity psychological intervention	CBT
Body Dysmorphic Disorder (BDD)	Step 2: Low-intensity psychological intervention	None
	Step 3: High-intensity psychological intervention	CBT

Note. Step 1 relates to the identification and assessment of CMDs within non-psychological primary-care settings, such as GP surgeries. Interventions can include active monitoring and collaborative care.

Appendix B

Table B1

Search Terms used to Conduct the Literature Search

	Search Terms and Operators
1	(psychotherap* OR (psycho* adj (aid* OR help* OR intervention* OR support* OR therap* OR training or treatment*))).mp
2	(psychoeducation* OR self-help OR counsel* OR Mindfulness OR Problem-solving OR Cognitive Behavioural Therap* OR CBT OR Behavioural Activation OR Cognitive Therap* OR CT OR stress manag* OR CCBT OR IPT OR interpersonal psychotherapy* OR DIT or EMDR OR dynamic interpersonal therap* rational emotive OR REBT OR schema OR solution focus*).mp
3	(computer* OR digital* adj2 (intervention* OR support* OR therap* OR training or treatment*)).mp
4	1 OR 2 OR 3
5	(common mental health disorder* OR common mental disorder*).mp
6	(Mood Disorders OR Depressive Disorder OR affective disorder* OR depressive disorder* OR depression* OR dysthymic disorder* OR dysthymia* OR affective disturbance* OR affective ill* OR mood disturbance*).mp
7	(Anxiety Disorders OR Neurotic Disorders OR Obsessive-Compulsive Disorder OR Panic Disorder OR Phobic Disorders OR Stress Disorders, Post-traumatic OR anxiety disorder* OR neurotic disorder* OR obsessive-compulsive disorder* OR panic disorder* OR phobic disorder* OR phobia* OR generalized anxiety disorder* OR generalised anxiety disorder* OR posttraumatic stress disorder* OR Body Dysmorphic Disorder).mp
8	(Trauma and Stressor Related Disorders OR Stress Disorders OR Traumatic OR Psychological Trauma OR Psychological Distress OR Stress OR Psychological OR trauma* OR stress disorder* OR psychological distress* OR emotional distress*).mp
9	5 OR 6 OR 7 OR 8
10	(Group-based trajector* OR latent class growth OR growth mixture OR GMM OR Response curve* OR Treatment expectanc*).mp
11	(trajector* adj2 symptom*).mp
12	10 OR 11
13	4 AND 9 AND 12

Appendix C

Table C1

Sample Patient Characteristics Pre- and Post-PSM with Goodness of Fit Statistics

Patient characteristic		Pre-PSM (n = 35547)						Post-PSM (n = 10268)					
		CBT (n = 29177)		Counselling (n = 6370)		Goodness of Fit		CBT (n = 4055)		Counselling (n = 6213)		Goodness of Fit	
		n	%	n	%	χ^2	p	n	%	n	%	χ^2	p
Age Band	16 – 24	5051	17.3%	403	6.3%	2219.00	< 0.001	345	8.5%	403	6.5%	16.56	0.346
	25 – 29	5758	19.7%	721	11.3%			528	13%	721	11.6%		
	30 – 34	4656	16%	777	12.2%			559	13.8%	772	12.4%		
	35 – 39	3461	11.9%	715	11.2%			485	12%	710	11.4%		
	40 – 44	2795	9.6%	647	10.2%			431	10.6%	639	10.3%		
	45 – 49	2438	8.4%	668	10.5%			417	10.3%	657	10.6%		
	50 – 54	1927	6.6%	647	10.2%			399	9.8%	631	10.2%		
	55 – 59	1446	5%	582	9.1%			338	8.3%	565	9.1%		
	60 – 64	779	2.7%	434	6.8%			230	5.7%	416	6.7%		
	65 – 69	399	1.4%	292	4.6%			131	3.2%	276	4.4%		
	70 – 74	224	0.8%	186	2.9%			90	2.2%	172	2.8%		
	75 – 79	140	0.5%	144	2.3%			60	1.5%	125	2%		
	80 – 84	64	0.2%	89	1.4%			23	0.6%	76	1.2%		
	85 – 89	30	0.1%	44	0.7%			11	0.3%	35	0.6%		
	90 – 94	4	0.0%	17	0.3%			4	0.1%	11	0.2%		
	Missing	5	0.0%	4	0.1%			4	0.1%	4	0.1%		
Gender	Not Reported	161	0.6%	101	1.6%	317.80	< 0.001	54	1.3%	91	1.5%	0.97	0.915
	Male	9634	33%	1491	23.4%			1041	25.7%	1479	23.8%		
	Female	19141	65.6%	4675	73.4%			2908	71.7%	4540	73.1%		
	Intermediate	30	0.1%	5	0.1%			3	0.1%	5	0.1%		

	Missing	211	0.7%	98	1.5%			49	1.2%	98	1.6%		
Ethnicity (ONS)	White	19015	65.2%	3782	59.4%	144.90	< 0.001	2458	60.6%	3686	59.3%	2.77	0.838
	Mixed	1646	5.6%	323	5.1%			212	5.2%	319	5.1%		
	Asian	2165	7.4%	589	9.2%			373	9.2%	574	9.2%		
	Black	2260	7.7%	513	8.1%			322	7.9%	499	8%		
	Chinese	214	0.7%	34	0.5%			16	0.4%	33	0.5%		
	Other	1010	3.5%	355	5.6%			197	4.9%	345	5.6%		
	Missing	2867	9.8%	774	12.2%			477	11.8%	757	12.2%		
LTC Declared	No	17654	60.5%	3041	47.7%	350.55	< 0.001	2071	51.1%	3000	48.3%	1.46	0.481
	Yes	5957	20.4%	1734	27.2%			1072	26.4%	1683	27.1%		
	Missing	5566	19.1%	1595	25%			912	22.5%	1530	24.6%		
Employment Status	Employed	16471	56.5%	3144	49.4%	1193.88	< 0.001	2101	51.8%	3112	50.1%	11.74	0.163
	Unemployed – Seeking Work	3687	12.6%	1037	16.3%			601	14.8%	1011	16.3%		
	Student	2590	8.9%	228	3.6%			192	4.7%	227	3.7%		
	Long-term Sick or Disabled	2074	7.1%	512	8%			365	9%	505	8.1%		
	Homemaker	1079	3.7%	296	4.6%			187	4.6%	287	4.6%		
	No Benefits and Not Seeking Work	1716	5.9%	199	3.1%			150	3.7%	198	3.2%		
	Unpaid Volunteer	113	0.4%	36	0.6%			22	0.5%	36	0.6%		
	Retired	948	3.2%	746	11.7%			344	8.5%	682	11%		
	Missing	499	1.7%	172	2.7%			93	2.3%	155	2.5%		
Psychotropic medication	Prescribed – Not Taking	4062	13.9%	466	7.3%	1101.14	< 0.001	382	9.4%	466	7.5%	83.39	< 0.001
	Prescribed – Taking	10756	36.9%	1879	29.5%			1389	34.3%	1877	30.2%		
	Not Prescribed	12618	43.2%	2926	45.9%			1876	46.3%	2888	46.5%		
	Missing	1741	6%	1099	17.3%			408	10.1%	982	15.8%		
Problem descriptor	Depression	7081	24.3%	1613	25.3%	2173.33	< 0.001	1161	28.6%	1602	25.8%	12.54	0.129
	Mixed Anxiety and Depression	1051	3.6%	202	3.2%			163	4%	202	3.3%		
	GAD	3062	10.5%	166	2.6%			152	3.7%	166	2.7%		

	OCD	1029	3.5%	1	0.0%			3	0.1%	1	0.0%		
	PTSD	1249	4.3%	49	0.8%			44	1.1%	49	0.8%		
	Phobic anxiety and Panic	3002	10.3%	20	0.3%			22	0.5%	20	0.3%		
	Unspecified Anxiety	145	0.5%	20	0.3%			23	0.6%	20	0.3%		
	SMI	30	0.1%	0	0.0%			0	0.00%	0	0.0%		
	Other	365	1.3%	266	4.2%			130	3.2%	250	4%		
	Missing	12163	41.7%	4033	63.3%			2357	58.1%	3903	62.8%		
	1	2691	9.2%	525	8.2%			350	8.6%	514	8.3%		
	2	7669	26.3%	1513	23.8%			947	23.4%	1472	23.7%		
	3	6386	21.9%	1114	17.5%			756	18.6%	1098	17.7%		
	4	3786	13%	839	13.2%			563	13.9%	811	13.1%		
	5	2656	9.1%	613	9.6%			372	9.2%	599	9.6%		
Index of Multiple Deprivation (IMD) Decile	6	2117	7.3%	557	8.7%	208.58	< 0.001	352	8.7%	545	8.8%	7.62	0.666
	7	1314	4.5%	437	6.9%			260	6.4%	423	6.8%		
	8	1263	4.3%	402	6.3%			238	5.9%	392	6.3%		
	9	471	1.6%	145	2.3%			95	2.3%	138	2.2%		
	10	108	0.4%	31	0.5%			19	0.5%	30	0.5%		
	Missing	716	2.5%	194	3%			103	2.5%	191	3.1%		
	1	7002	24%	1966	30.9%			1389	34.3%	1950	31.4%		
Trust	2	15597	53.5%	4195	65.9%	1300.03	< 0.001	2445	60.3%	5054	65.3%	46.52	< 0.001
	3	3793	13%	16	0.3%			19	0.5%	16	0.3%		
	4	2785	9.5%	193	3%			202	5%	193	3.1%		
	Low	1983	6.8%	357	5.6%			210	5.2%	343	5.5%		
	Mild	4852	16.6%	1071	16.8%			605	14.9%	1035	16.7%		
PHQ Severity Band	Moderate	7214	24.7%	1733	27.2%	28.17	< 0.001	1049	25.9%	1688	27.2%	11.44	0.043
	Moderate Severe	7601	26.1%	1638	25.7%			1082	26.7%	1603	25.8%		
	Severe	7339	25.2%	1539	24.2%			1079	26.6%	1512	24.3%		
	Missing	188	0.6%	32	0.5%			30	0.7%	32	0.5%		
GAD Severity Band	Low	1658	5.7%	549	8.6%	207.33	< 0.001	297	7.3%	517	8.3%	8.25	0.083
	Mild	5265	18%	1435	22.5%			853	21%	1392	22.4%		

	Moderate	8361	28.7%	1880	29.5%			1125	27.7%	1831	29.5%		
	Severe	13690	46.9%	2471	38.8%			1748	43.1%	2438	39.2%		
	Missing	203	0.7%	35	0.5%			32	0.8%	35	0.6%		
WSAS Severity Band	Mild	3465	11.9%	958	15%	55.35	< 0.001	510	12.6%	908	14.6%	7.02	0.071
	Moderate	7783	26.7%	1707	26.8%			1062	26.2%	1662	26.8%		
	Severe	9592	32.9%	2050	32.2%			1362	33.6%	2014	32.4%		
	Missing	8337	28.6%	1655	26.0%			1121	27.6%	1629	26.2%		

Appendix D

Table D1

LGCA Statistics for Model Fit

Model	PHQ-9					GAD-7				
	χ2 (df)	CFI	TLI	RMSEA	SRMR	χ2 (df)	CFI	TLI	RMSEA	SRMR
Linear	5033.17 (145)	0.919	0.924	0.057	0.128	4276.94 (145)	0.926	0.931	0.053	0.125
Quadratic	2859.74 (140)	0.955	0.956	0.043	0.056	2475.99 (140)	0.958	0.960	0.040	0.054

Note. Models include treatment type (CBT and Counselling) as a predictor. CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = The Root Mean Square Error of Approximation; SRMR = Standardised Root Mean Residual.

Table D2

LGCA Linear Model Parameter Statistics

Parameter	PHQ-9		GAD-7	
	Mean	95% CI	Mean	95% CI
Intercept	13.64 (0.079)	13.46 – 13.79	11.87 (0.067)	11.739 – 12.001
Linear	-0.479 (0.009)	-0.474 – -0.484	-0.394 (0.008)	-0.410 – -0.378

Note. Standard error values presented in parentheses. CI = Confidence Interval.

Appendix F

Table F1

GMM Statistics for k-1 Model Fit

PHQ-9						
k model	AIC	BIC	Adj-BIC	VLMR-LRT *p*	Entropy	Class Membership %
k = 2	546647.226	546857.092	546764.935	< 0.001	0.679	78/22
k = 3	546289.854	546528.668	546423.798	< 0.001	0.582	19/57/24
***k* = 4**	**545997.565**	**546265.326**	**546147.745**	**< 0.001**	**0.631**	**7/8/60/25**
k = 5	545735.292	546032.000	545901.708	0.2309	0.634	7/30/4/13/46
GAD-7						
k model	AIC	BIC	Adj-BIC	VLMR-LRT *p*	Entropy	Class Membership %
k = 2	529760.083	529969.950	529877.792	< 0.001	0.649	73/27
k = 3	529176.744	529415.558	529310.688	< 0.001	0.599	48/29/22
***k* = 4**	**528837.930**	**529105.691**	**528988.110**	**< 0.001**	**0.607**	**13/9/48/31**
k = 5	528619.266	528915.974	528785.682	0.2907	0.625	26/20/28/7/9

Note. *k* = class number. Chosen model highlighted in bold. AIC = The Akaike Information Criterion; BIC = Bayesian Information Criterion; VLMR-LRT = Vuong-Lo-Mendell-Rubin Likelihood Ratio Test.

Table F2

GMM Class Model Parameter Statistics

Class	Intercept		Linear		Quadratic	
	Mean (*SE*)	95% CI	Mean (*SE*)	95% CI	Mean (*SE*)	95% CI
PHQ-9 GMM Classes						
Class 1	19.098 (0.392)	18.330 – 19.866	-3.053 (0.328)	-3.696 – -2.410	0.162 (0.026)	0.111 – 0.213
Class 2	19.010 (0.013)	18.985 – 19.035	-0.578 (0.022)	-0.621 – -0.535	-0.042 (0.022	-0.085 – 0.001
Class 3	10.776 (0.240)	10.306 – 11.247	-0.793 (0.054)	-0.900 – -0.687	0.034 (0.004)	0.026 – 0.042
Class 4	18.583 (0.195)	18.201 – 18.965	-0.031 (0.064)	-0.156 – -0.094	-0.005 (0.05)	-0.103 – 0.093
GAD-7 GMM Classes						
Class 1	15.785 (0.292)	15.273 – 16.357	-2.233 (0.219)	-2.662 – -1.804	0.112 (0.017)	0.079 – 0.145
Class 2	16.025 (0.210)	15.613 – 16.437	-0.302 (0.180)	-0.655 – 0.051	-0.049 (0.019)	-0.086 – -0.012
Class 3	8.270 (0.144)	7.988 – 8.552	-0.535 (0.040)	-0.613 – -0.457	0.022 (0.003)	0.016 – 0.028
Class 4	8.163 (0.391)	7.397 – 8.929	0.441 (0.031)	0.380 –0.502	0.002 (0.000)	0.002 – 0.002

Note. SE = Standard Error; CI = Confidence Interval.