



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/238164/>

Version: Published Version

Article:

Kamardina, A., Saunders, R. and Broglia, E. (2026) The impact of initial depression and anxiety levels on symptom severity change during psychological treatment for common mental disorders: a systematic review. *Journal of Mental Health*, 35 (3). pp. 354-374. ISSN: 0963-8237

<https://doi.org/10.1080/09638237.2026.2622085>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



The impact of initial depression and anxiety levels on symptom severity change during psychological treatment for common mental disorders: a systematic review

Anna Kamardina, Rob Saunders & Emma Broglia

To cite this article: Anna Kamardina, Rob Saunders & Emma Broglia (05 Feb 2026): The impact of initial depression and anxiety levels on symptom severity change during psychological treatment for common mental disorders: a systematic review, Journal of Mental Health, DOI: [10.1080/09638237.2026.2622085](https://doi.org/10.1080/09638237.2026.2622085)

To link to this article: <https://doi.org/10.1080/09638237.2026.2622085>



© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 05 Feb 2026.



[Submit your article to this journal](#)



Article views: 215



[View related articles](#)



[View Crossmark data](#)

The impact of initial depression and anxiety levels on symptom severity change during psychological treatment for common mental disorders: a systematic review

Anna Kamardina^a , Rob Saunders^b  and Emma Broglia^c 

^aStudent Mental Health, Counselling & Therapies Service, University of Sheffield, Sheffield, UK; ^bResearch Department of Clinical, Educational and Health Psychology, CORE Data Lab, Centre for Outcomes Research and Effectiveness (CORE), University College London, London, UK; ^cDepartment of Psychology, University of Sheffield, Sheffield, UK

ABSTRACT

Background: Depression and anxiety are highly comorbid conditions, which may impact treatment response.

Aims: We intended to examine how depression and anxiety symptoms change during psychological interventions for adults and how the initial severity is associated with distinct trajectory classes.

Methods: We identified papers examining longitudinal changes in depression and anxiety. Trajectory classes were grouped into responders, improvers and non-improvers. Narrative synthesis, alongside meta-analysis, was applied to explore the number and nature of classes, their associations with study characteristics and predictors of trajectory class membership. This study was pre-registered on International Prospective Register of Systematic Reviews (PROSPERO) (CRD42023393155), and the authors followed Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidance.

Results: We identified 18 papers, including 22 models with two to six classes. Responding and non-improving classes prevailed. Meta-regression revealed no characteristics associated with the proportion of these classes. Initial depression and anxiety levels predicted trajectory class membership for each other, but not always non-response. Unemployment, suicidality, social functioning impairment, treatment type and medication use predicted following non-responding trajectory.

Conclusion: Our findings highlight the complex relationship between initial anxiety and depression severity and their changes during psychological interventions. They provide practitioners with insights into factors associated with non-response patterns of symptom change.

ARTICLE HISTORY

Received 13 March 2025
Revised 5 November 2025
Accepted 8 January 2026

KEYWORDS

depression; anxiety;
psychological
intervention; trajectory;
review

Introduction

Common mental disorders (CMDs), such as depression and anxiety, are estimated to affect 29% of the population globally at some point (Steel et al., 2014). In addition to individual suffering, they are also linked to a considerable burden on healthcare systems and wider society (McCrone, 2008). Anxiety and depression are also highly comorbid conditions. Almost 2/3 of individuals diagnosed with generalised anxiety disorder (GAD) develop a mood disorder over their lifetime (Ruscio et al., 2017). The WHO World Mental Health Survey found that almost half of people with lifelong major depressive disorder

(MDD) were diagnosed with at least one lifelong anxiety disorder. Of those who had 12-month MDD, 52% reported comorbidity with a lifetime anxiety disorder, and 42% reported comorbidity with a 12-month anxiety disorder (Kessler et al., 2015).

Numerous studies support the effectiveness of psychological interventions (e.g. Cuijpers et al., 2014, 2023), but over half of patients still do not respond to treatment (Cuijpers et al., 2024; Moller et al., 2019). In addition, the influence of comorbidity on treatment outcomes remains disputable. Previous research suggests that comorbid anxiety negatively impacts depression treatment outcomes (Buckman et al., 2021; Tunvirachaisakul et al., 2018). However, the former

CONTACT Anna Kamardina  a.kamardina@sheffield.ac.uk  Student Mental Health, Counselling & Therapies Service, University of Sheffield, Sheffield, UK

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/09638237.2026.2622085>.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

meta-analysis highlighted the importance of adjusting for depression severity, which resulted in some anxiety-related clinical variables becoming non-significant predictors of depression treatment outcomes. The latter study was limited by the age group, treatment type and the number of studies. When examining the depression comorbidity impact on anxiety treatment outcomes, mixed results were reported depending on the psychological intervention type (Newman et al., 2019; Wolitzky-Taylor et al., 2012).

The observed differential response to psychological interventions led to the recent movement towards personalised treatment approaches in mental health, focusing not only on the diagnoses but also on the patient's profile (Delgadillo & Gonzalez Salas Duhne, 2020; Saunders et al., 2016). Interventions tailored to individual patients were found to be more effective than standardised treatments (Nye et al., 2023). However, these studies used baseline characteristics to predict treatment eventual outcomes but not patterns of change.

Analysing within-treatment data can help practitioners better address their patients' needs and improve outcomes. Growth mixture modelling (GMM, Muthén, 2004, 2001) is one approach to exploring this. It identifies trajectory classes based on individual symptom change patterns, allowing practitioners to track interindividual differences more accurately. Moreover, GMM helps identify potential predictors of class allocation, enabling practitioners to recognise risk factors for non-improving trajectories.

The existing reviews on CMD trajectory classes partially addressed the effects of comorbidity. Thus, in a review on post-traumatic stress disorder (PTSD) Dewar et al. (2020) analysed 11 studies and identified comorbid depression and anxiety symptoms as strong predictors of PTSD trajectory class allocation. This study intends to extend these findings, focusing on the interaction of anxiety and depressive symptoms during psychological interventions and addressing the following research questions:

1. Which forms of trajectory are identified for anxiety and depression symptom change (in studies that measure both symptoms), and how do the number and nature of trajectories relate to study characteristics?
2. How are initial levels of anxiety and depression symptom severity associated with likely trajectory class followed?
3. Are individuals most likely to follow similar trajectories of depression and anxiety symptom change during psychological interventions?

4. Which other factors are associated with likely trajectory classes reported for anxiety and depression?
5. How does modelling both symptoms in the same model affect the results of trajectory analysis?

Methods

The protocol for this systematic review was registered with the International Prospective Register of Systematic Reviews (PROSPERO, CRD42023393155). The review follows the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Page et al., 2021).

Search method

The literature search, covering peer-reviewed journal papers in English, was conducted through Scopus, PubMed, PsycINFO, and Cochrane Library up to 10 June 2023. A manual citation search through the final study list and existing systematic reviews aimed to identify relevant papers. [Supplemental Material A](#) reveals the search terms.

Inclusion and exclusion criteria

We included studies with participants over 18 years old with anxiety disorders and/or depression seeking psychological intervention for CMDs. The condition was defined by meeting recognised diagnostic criteria or exceeding a clinical threshold on validated screening or symptom severity measures. In studies using the presenting issue concept (where clients sought help for anxiety or depressive symptoms), clinical cut-off requirements were not always applied and participants with severe mental health conditions were possible. Studies of evidence-based psychological interventions for CMD (e.g. cognitive behavioural therapy (CBT), psychodynamic therapy (PDT), counselling, guided self-help interventions, etc.) delivered in community and outpatient settings in any format (e.g. in-person or online, individual or group format) were included.

In addition, we included studies that identified distinct depression and/or anxiety symptom trajectories during treatment, even if only one class was extracted. This encompassed studies employing GMM, group-based trajectory modelling, and similar approaches. Eligible studies explored at least three measurements and used validated measures of anxiety and depression. If only one condition (e.g. depression) was modelled, we included only those studies where

at least one measurement of the other condition (e.g. anxiety) was analysed as a potential predictor of the modelled condition trajectory.

We excluded studies that focused on PTSD, as the condition was excluded from the DSM-5 list of anxiety disorders by the American Psychiatric Association (2013) and because previous research has explored trajectories for PTSD (Dewar et al., 2020). We also excluded studies that focused on severe mental health conditions (e.g. schizophrenia, personality disorders, and bipolar disorder), addictive disorders, relapse prevention, and studies evaluating medication interventions only.

No restrictions on the publication date or study type were applied.

Data extraction

After the database search, duplicates were removed, and titles and abstracts were screened for relevance. Eligible studies were then assessed in full, with unavailable full texts requested from authors.

The final list of eligible studies informed data extraction, including sample size and type, sample condition, clinical measures used, measurement frequency, intervention type, its format and duration, trajectory analysis procedure, the way of modelling anxiety and depression symptom changes, number and nature of trajectory classes reported, class allocation, predictors of class allocation, and the original data source (in case of secondary data).

Quality assessment

The reporting quality of the studies was measured with the GRoLTS-Checklist (Guidelines for Reporting on Latent Trajectory Studies, van de Schoot et al., 2017). We did not exclude studies based on their GRoLTS score, as currently there are no thresholds distinguishing low and high scores.

Four papers (22%) were randomly selected and assessed by an independent rater. The agreement between the rater and the author (AK) was evaluated by the percent of agreement and Cohen's kappa; discrepancies were discussed, and consensus was agreed.

Analysis

Research question #1. Number and nature of anxiety and depression trajectory classes and their association with study characteristics

To address this research question, we conducted narrative synthesis and quantitative methods where possible.

Trajectory class classification into “responders”, “improvers”, and “non-improvers” was conducted based on the mean predicted scores using a reliable change index (RCI) and clinically significant change approach. A class was labelled as “non-improver” if by the end of the modelled treatment period no improvement at RCI was reported, as “improver” if RCI was achieved and as “responder” if in addition to achieving RCI, the “caseness” threshold was crossed. For the study where two conditions were modelled together, we labelled classes based on meeting the criteria for at least one condition. Where RCI and “caseness” thresholds were not published, we performed our calculations following instructions of Jacobson and Truax (1991). [Supplemental Material B](#) reveals the details of the thresholds applied.

The relationship between study characteristics and the number of classes was examined via Pearson's correlation and linear regression modelling. Random effects meta-analysis on the proportion of non-improving class and responding class was conducted where classification was possible. When analysing both conditions together, for those studies where both depression and anxiety were modelled, we averaged proportions of each class across two conditions. To examine potential sources of heterogeneity, we conducted meta-regression analysis using moderators available, including average sample age, sample female proportion, the application of a clinical threshold to filter participants (yes/no), intervention type (CBT-based/others), the intervention duration (in weeks), the type of measure (self-report/clinician-administered), and follow-up inclusion (yes/no).

The quantitative analysis was conducted in R Statistical Software (v.4.2.1, R Core Team, 2022). We ran a meta-analysis with meta R package (Balduzzi et al., 2019), using the inverse variance method, REML estimation, and normal approximation confidence interval for individual studies; logit transformation of the class proportions and Hartung–Knapp adjustment (Knapp & Hartung, 2003) to calculate the confidence intervals around the pooled effect were applied. [Supplemental Material C](#) reveals the code.

Research questions #2–5

A narrative synthesis was applied to address the remaining research questions.

All the studies were allocated into three groups: (i) studies modelling one of the two conditions (either anxiety or depression) and measuring the other one at baseline; (ii) studies modelling both anxiety and depression for the same sample but within separate models; (iii) studies modelling anxiety and depression symptom changes within a single model.

Research question #2. The association of initial anxiety and depression symptom severity with class membership

A narrative synthesis was applied across all the studies to assess the association between class membership and the initial severity of the condition modelled. Studies from the first two groups above were used to examine the association between class membership and the comorbid condition severity (i.e. anxiety initial severity for depression class membership and vice versa).

Research question #3. Comparison of individual trajectories across the two conditions

To address this question, studies modelling both conditions (i.e. the second and third study groups) were used.

Research question #4. Other predictors of class membership

We summed up the results across all the studies where authors performed analysis on potential predictors for anxiety and/or depression class membership.

Research question #5. The effect of modelling the two conditions together vs. modelling them separately

We aimed to address this question by comparing the second and third study groups. However, the third

group had just one study, leaving the question unanswered in this review.

Results

The initial search identified 1,692 papers (895 after removing duplicates), out of which 65 papers were selected for further full-text screening, leading to 18 papers in the final list. Manual citation search provided additional 14 papers, none of which satisfied the inclusion criteria. Figure 1 presents the PRISMA flow diagram, with the included studies presented in Table 1.

Eleven (61% of all included papers) papers used secondary data, and seven studies (39%) used routinely collected data from psychological services or licensed therapists working independently. No overlapping datasets were identified.

In eight papers (44%), authors modelled trajectories of depressive symptoms. Five studies (28%) modelled trajectories of anxiety (Joesch et al., 2013; Lukaschek et al., 2019) and anxiety disorders: panic disorder (Lutz et al., 2014), obsessive-compulsive disorder (H. Kim et al., 2023), and social anxiety disorder (Altmann et al., 2020). Four studies (22%)

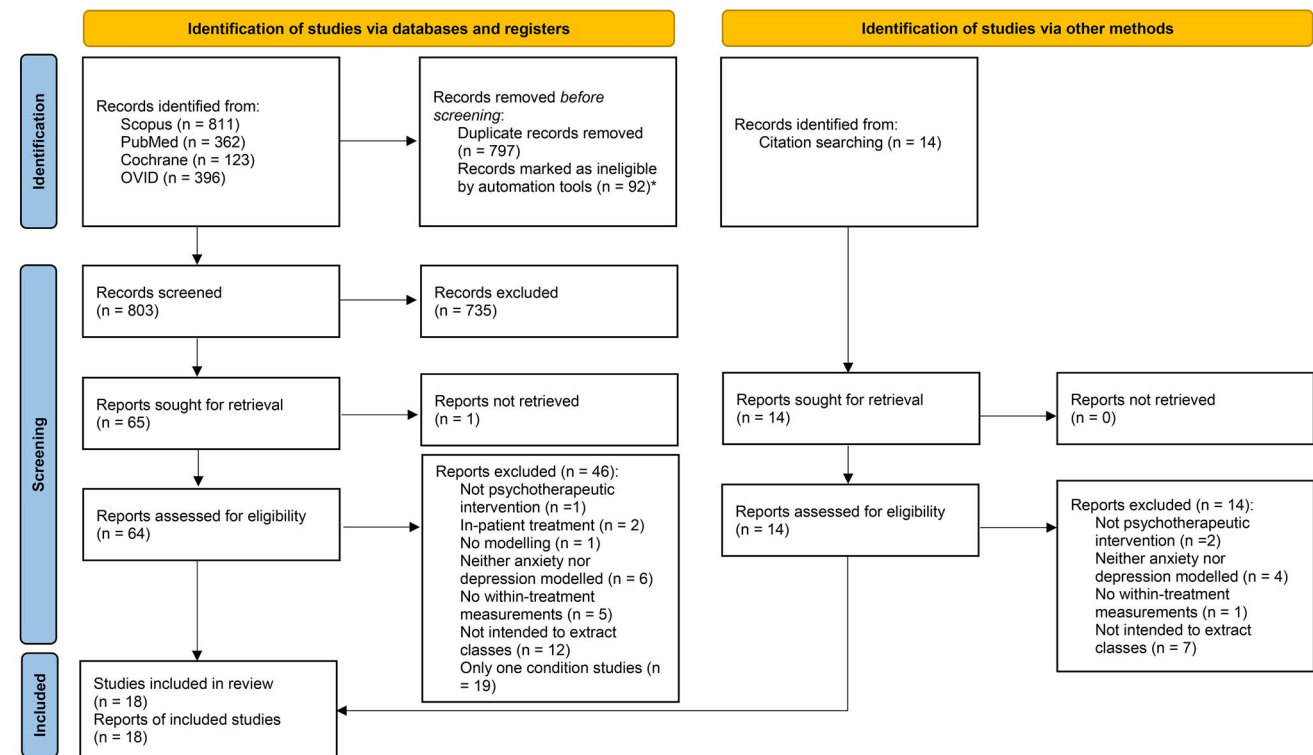


Figure 1. Flow diagram. *Automated title search was conducted for "bipolar", "PTSD", and "posttraumatic stress" to remove the relevant papers before manual screening. Source: Page et al. (2021), licensed under CC BY 4.0.

Table 1. Studies included in the review and their key characteristics (in alphabetical order).

#	Study	Sample condition	Clinical sample ^b	Clinical exclusion criteria ^c	Sample size	Average age (SD)	Prop. of females	Ethnicity	Intervention type	Format of delivery	Intervention duration	Type of study	GROLTS score
1	Altmann et al. (2020)	Social anxiety	Yes	Yes	357	34.6 (7.9)	55%	–	Short-term CBT and PDT for SAD	In-person	Up to 25 sessions	RCT	10
2	Batterham et al. (2017)	Depression with insomnia	No	Yes	574	43 (12)	74%	–	CBT-I (an insomnia program)	Online	8 weeks	RCT	7
3	Cuijpers et al. (2005)	Depression	Yes	Yes	425	36 (10)	61%	–	CBT and TAU	In-person	Mean duration 224 days (SD = 202), 10 sessions in average (SD = 11)	Randomised trial	11
4	Hull et al. (2020)	Depression and anxiety	Yes	Yes	10,718	Range 18–65	79%	–	Mixture (therapists are mainly CBT oriented, 25% psychodynamic or relational)	Tele-medicine	12 weeks with messaging up to 5 days a week	Routinely collected	13
5	Jankowski et al., (2019)	Clients receiving long-term outpatient treatment	No	No	118	41 (12)	53%	White 81%	Long-term CRP	In-person	Long-term	Routinely collected	13
6	Joesch et al. (2013)	Anxiety disorder (with high comorbidity with depression)	Yes	No	482	43 (13)	70%	Diverse race/ethnicity	CALM (CBT, anti-anxiety medication, or both)	Online	10–12 weeks	RCT	8
7	Kelley et al. (2018)	Depression	Yes	No	344	40 (12)	57%	White 49% Hispanic 32%	CBT, medication	In-person	12 weeks/16 sessions	Randomised study	10
8	H. Kim et al. (2023)	OCD	Yes	Yes	334	33	46%	White 93%	Exposure and response prevention	In-person	8 weeks/17 sessions	Clinical trials	8
9	Lukaschek et al. (2019)	Panic disorder	Yes	Yes	176	45 (14)	72%	–	CBT (psychoeducation and CBT-oriented exposure exercises)	In-person	23 weeks	RCT	6
10	Lutz et al. (2014)	Panic disorder	Yes	Yes	326	37 (12)	64%	White 88%	CBT (a modified version of Panic Control Treatment)	In-person	11-session administered on a weekly (biweekly for the last two sessions) basis	Randomised trial	11
11	McDevitt-Petrovic et al. (2020)	People seeking help at IAPT	No	No	250	38 (14)	60%	–	Low intensity CBT	In-person	In the range of 1–11 sessions	Routinely collected	9
12	Nordberg et al. (2014)	Depression	Yes	No	147	39 (12)	71%	Caucasian 86%	CBT, PDT	In-person	Minimum 15 sessions	Routinely collected	12
13	Saunders et al. (2019)	People seeking help at IAPT	Yes	No	4394	38 (13)	63%	White 67%	High-intensity psychological therapy for CMD (CBT, counselling, and IPT)	In-person	Up to 12 sessions included in analysis	Routinely collected	11
14	Siddique et al. (2012)	Depression	Yes	Yes	267	29	100%	Latina 50% Black 44%	CBT, TAU, and medication	Group or individual in-person sessions	8 weekly sessions	RCT	10
15	Skelton et al. (2023)	CMD	No	No	16,258	38 (13)	68%	White 60.2%	High-intensity psychological therapy for CMD (disorder-specific CBT, counselling, IPT)	Supposed to be in-person	The number of sessions ranged in 3–55	Routinely collected	18

(Continued)

Table 1. Continued.

#	Study	Sample condition	Clinical sample ^b	Clinical exclusion criteria ^c	Sample size	Average age (SD)	Prop. of females	Ethnicity	Intervention type	Format of delivery	Intervention duration	Type of study	GROLTS score
16	Smolenski et al. (2017)	Depression	Yes	No	121	Range 19–65	18%	–	Behavioural activation for depression	Via videoconference or in-person	8 weeks with 50-min weekly sessions	RCT	6 ^d
17	Stulz et al. (2010)	Chronic depression	Yes	Yes	504	43 (10)	64%	White 92%	Medication, CBASP, and their combination	Supposed to be in-person	12 weeks/16 sessions	RCT	11
18	Vetter et al. (2022)	Depression	Yes	No	239	41	57%	–	Evidence-based PDT and CBT	In-person, individual and group sessions	Determined individually	Routinely collected	10
Total					36,034								
Average					2002	37 ^a	70% ^a						10

CALM: coordinated anxiety learning and management; CBASP: cognitive behavioural analysis system of psychotherapy; CBT: cognitive behavioural therapy; CRP: contemporary relational psychotherapy; IPT: interpersonal psychotherapy; OCD: obsessive compulsive disorder; PDT: psychodynamic therapy; RCT: randomized control trial.

^aAverage weighted by sample size.

^bClinical sample marked as yes if the authors reported that participants were initially screened for having clinical scores on the relevant measure or diagnosed with the modelled condition.

^cClinical exclusion criteria marked as yes if the authors reported exclusion of participants with severe mental conditions, usually included bipolar disorder, schizophrenia spectrum, psychotic disorder, or psychotic features, substance/alcohol use disorder, suicidal thoughts and/or behaviour.

^dAs a conditional two-class model did not converge and thus one-class solution was reported as the final model for depression, the paper reporting quality was assessed based on reporting trajectory modelling for hopelessness (the other construct modelled in the study and reported to have two trajectory classes). The number of classes extracted for depression was coded as two, as the two-class solution did not converge only after adding covariates but was one that better fitted the data.

modelled both anxiety and depression for the same samples within separate models (Jankowski et al., 2019; McDevitt-Petrovic et al., 2020; Saunders et al., 2019; Skelton et al., 2023) and one study modelled two conditions in a single model (Hull et al., 2020).

Sample sizes varied widely, from 118 to 16,258, with a mean of 2,002 and a median of 339. Average participant age ranged from 29.0 to 45.3 years, while the proportion of female participants ranged from 18% to 100%.

In 14 studies (78%), clinical inclusion criteria were applied (a condition was diagnosed and confirmed via a clinical interview and/or by a clinical threshold of a respective measurement tool). In one study (Batterham et al., 2017), the criterion was applied to insomnia rather than depression (participants were scored at 4–20 by PHQ-9, thus, the study was marked as one with mixed population). In three studies (Jankowski et al., 2019; McDevitt-Petrovic et al., 2020; Skelton et al., 2023), authors conducted no participant clinical screening and all routinely collected data were accepted for analysis.

Interventions were grouped into four types: mixed ($k = 7$), CBT-based ($k = 6$), mixed with medication ($k = 4$) and psychodynamic ($k = 1$). In three studies (Batterham et al., 2017; Hull et al., 2020; Joesch et al., 2013), interventions were delivered remotely; in one study (Smolenski et al., 2017) in-person and online formats were compared. All other cases involved in-person interventions, including two studies (Skelton et al., 2023; Stulz et al., 2010) without delivery format specification, but their setting descriptions indicated face-to-face delivery.

Most of the studies were performed in the USA ($k = 11$), three in the UK, two in Germany, one in Australia, one in Switzerland and one in the Netherlands. All the papers were published after 2004.

Quality assessment

The percent of agreement between two raters comprised 0.87, and Cohen's kappa ratio was 0.72 ($z = 6.63$, $p < 0.001$), which is assessed as moderate to substantial (McHugh, 2012).

The average total score on the GROLTS-checklist comprised 10.2 of 21 with a range of 6–18. The total GROLTS score was associated with study sample size (Pearson's $r = 0.73$ [95%CI: 0.40; 0.89], $p < 0.001$); no association was found with the year of publication.

The most frequently reported items were covariate analysis, model comparison tools, software used and plot of mean trajectories of the final model. The least frequently reported were plots of trajectories for all

models considered and of individual trajectories and starting values. No syntax files were found. [Supplemental Material D](#) reveals more details.

Trajectory analysis

Research question #1. Number of trajectory classes

The number of trajectory classes reported varied from two to six (mean = 3.25). Only one study (Hull et al., 2020) modelled both conditions in a single model and reported six classes, which was treated as an outlier because of the different methodology. Without it, the mean number of classes was 3.1.

Out of 21 models from 17 studies (excluding the outlier), 12 modelled depression symptoms and nine anxiety symptoms. The average number of classes extracted for anxiety was 3.3, which was slightly higher than for depression (2.9), but this difference was not significant ($F = 1.00$, $p = 0.329$).

The number of classes was associated with the sample size (Pearson's $r = 0.57$ [95%CI: 0.14; 0.82], $p = 0.014$), but after excluding the outlier, the coefficient became non-significant even before Bonferroni's correction ($r = 0.38$ [95%CI: -0.12; 0.73], $p = 0.132$). [Supplemental Material E](#) reveals associations between the number of classes and study/sample characteristics.

We also checked whether the number of classes was associated with the sample type (clinical/mixed), follow-up inclusion (yes/no), the intervention type (CBT-based/other) and its delivery method (in-person/remotely). F -statistics for none of these reached statistical significance. When controlling for the sample size, the type of the sample was the only statistically significant predictor, adding 32% of explained variance in the number of classes in addition to that explained by the sample size ($b = 1.51$, $SE = 0.41$, $p = 0.002$, $R^2 = 64.5\%$, $F(2,15) = 13.63$, $p < 0.001$). Removing the outlier (Hull et al., 2020) did not affect these results. The mean number of classes was 3.5 (3.3 without the outlier) in studies with the clinical population vs. 2.5 in studies with the mixed population.

Research question #1. Nature of trajectory classes

Sixty-five trajectory classes from 20 models (16 studies) were classified into responders (30 classes), improvers (7) and non-improvers (28). Non-improvers included one deteriorating class reported by Lutz et al. (2014), modelling the first five sessions (the shortest period among the studies). [Table 2](#) reveals the classification results.

The proportion of non-improving class varied considerably across studies, from 0 to 86% (Nordberg et al., 2014). One study (Stulz et al., 2010) reported no non-improvers, possibly due to excluding participants with previous non-response to treatment.

Research question #1. Meta-analysis and meta-regression

Meta-analysis was conducted for 15 studies, where class classification was possible and at least one non-improver class was extracted.

Non-improving class proportion. The pooled proportion of the non-improving class was 36.1% [95%CI: 20.9; 54.8%], although high heterogeneity was observed: $Q(14) = 2133$, $p < 0.0001$, $I^2 = 99.3\%$, 3–92% prediction interval for the pooled effect. [Figure 2](#) presents the forest plot. Meta-regression analysis identified no predictors of the non-improver class proportion.

The pooled effect was similar when meta-analysis was conducted separately for studies modelling either depression or anxiety alone. For depression 36.4% [95%CI: 13.9; 67.0%] non-improving class proportion was reported with high heterogeneity: $Q(8) = 1156$, $p < 0.0001$, $I^2 = 99.3\%$. For anxiety, it was 40.9% [95%CI: 21.7; 63.4%] non-improving class proportion with also high heterogeneity: $Q(8) = 1568$, $p < 0.0001$, $I^2 = 99.5\%$. No predictors were identified for the non-improving proportion across the two conditions.

Responder class proportion. The pooled proportion of the responder class comprised 34.3% [95%CI: 11.7; 67.2%]. High heterogeneity was indicated too: $Q(14) = 2371$, $p < 0.0001$, $I^2 = 99.4\%$, 0–99% prediction interval. [Supplemental Material F](#) reveals the forest plot. In unadjusted models, the proportion of the responder class was negatively associated with the proportion of females ($b = -11.9$, $p = 0.020$) and the follow-up inclusion ($b = -3.5$, $p = 0.036$, i.e. it was higher when only the treatment period was modelled), but it was no longer significant when adjusted for other predictors.

When analysing separately depression and anxiety models, the pooled responder class proportion was higher for anxiety: 50.9% [95%CI: 27.6; 73.8%] vs. 24.3% [95%CI: 2.8; 78.4%] reported for depression; although overlapping CIs and high heterogeneity were reported (anxiety: $Q(8) = 1467$, $p < 0.0001$, $I^2 = 99.5\%$; depression: $Q(8) = 710$, $p < 0.0001$, $I^2 = 98.9\%$). No predictors were reported for the responding class proportion based on studies modelling anxiety; for depression models, meta-regression provided the same results as for all studies: the proportion of females ($b = -18.7$, $p = 0.011$) and the follow-up inclusion

Table 2. Results of classification of trajectory classes (ordered according to the frequency of the measure used).

Study	Sample size	Measurement frequency	Measure used	Modelling method	Period modelled	No. of classes	Responders	Improvers	Non-improvers
<i>Studies modelling depression</i>									
McDevitt-Petrovic et al. (2020)	250	At every session	PHQ-9	GMM	First six sessions only	2	Improvers (87.2%)		Non-improvers (12.8%)
Saunders et al. (2019)	4394	At every session	PHQ-9	LCGA	Within treatment only	4	Moderate-to-severe pre-treatment, slow initial response, large response later in treatment (13.5%); moderate-to-severe pre-treatment, rapid early improvement, levelling out (14.0%)		Moderate severity pre-treatment, early initial response, levelling out (44.5%); moderate-to-severe pre-treatment, limited or no response to treatment (28%)
Skelton et al. (2023)	16,258	At every session	PHQ-9	GMM	Within treatment only	4	Initially moderate-severe symptoms and fast improvement (16.4%); initially moderate-severe symptoms and gradual improvement (17.6%)		Initially moderate-severe symptoms and no change (13.5%); Initially mild symptoms and minimal improvement (52.5%)
Batterham et al. (2017)	574	At baseline, at 2/4/6/8 weeks, post-intervention, and at 6/12/18 m follow-up	PHQ-9 without sleep item	LGMM	Within treatment and follow-up	2		Improving (95 %)	Stable/deteriorating (5%)
Kelley et al. (2018)	344	At baseline, weekly through week 6, and at weeks 8/10/12	HDRS-17	LCA	Within treatment only	4	Early remission (3 weeks) 23.8%; Gradual remission (12 weeks) – 39.8%; minimal response – 26.5%		No response – 9.9%
Siddique et al. (2012)	267	At baseline, monthly for 6 months, and at months 8/10/12	HDRS-17	GMM	Within treatment and follow-up	2		Moderate class (64.5%)	Severe class (35.5%)
Vetter et al. (2022)	239	At baseline, after six weeks, and post-treatment	HDRS-17	LGMM	Within treatment only	3	Responding from moderate severity (72.8%)	Responding from severe depression (7.5%)	Non-responding class, moderately to severely depressed at baseline (19.7%)
Jankowski et al. (2019)	118	No regular frequency reported (at multiple points during and at the end of treatment)	TOP depression scale	LGM	Within treatment only	2	Class 1 (36.5%) – higher initial depression/reliable improvement		Class 2 (63.5%) – lower initial depression scores/not significant trajectory, no improvement
Nordberg et al. (2014)	147	At baseline, at sessions 7 and 15, and every 15 sessions thereafter	TOP depression scale	GMM	Within treatment only	3	High symptom rapid responders (13.6%)		Not-responders with low symptoms (58.5%); not-responders with high symptoms (27.9%)
Stulz et al. (2010)	504	At baseline, at weeks 1–4/6/8/10/12 and post-treatment	HDRS-24	GMM	Within treatment only	3	Severe depression (20.2%); moderate depression (58.3%)	Low depression (21.4%)	
Cuijpers et al. (2005)	425	At baseline, then every 3 months until 7 measurements were taken	SCL-90	GMM	Within treatment and follow-up	4	No information available for classification. All 4 classes improved with approximately the same rate of improvement but different intercepts		
Smolenski et al. (2017)	121		BDI-II	GMM	Within treatment and follow-up	2	Only one improving class (100%) reported based on the small sample (the two-class model was better but it was not stable with covariates and thus not reported)		

(Continued)

Table 2. Continued.

Study	Sample size	Measurement frequency	Measure used	Modelling method	Period modelled	No. of classes	Responders	Improvers	Non-improvers
Mean	1970					2.9	33% ^a	3% ^a	64% ^a
Median	306					3.0			
<i>Studies modelling anxiety and anxiety disorders</i>									
McDevitt-Petrovic et al. (2020)	250	At every session	GAD-7	GMM	First six sessions only	2	Improvers (82.8%)		Non-improvers (17.2%)
Saunders et al. (2019)	4394	At every session	GAD-7	LCGA	Within treatment only	5	Severe pre-treatment, slow initial response, large response later in treatment (11.6%); moderate-to-severe pre-treatment, early initial response, levelling out (28.6%); mild pre-treatment, early initial response, levelling out (17.4%)		Moderate-to-severe pre-treatment, limited or no response to treatment (21.1%); severe pre-treatment, limited or no response to treatment (21.4%)
Skelton et al. (2023)	16,258	At every session	GAD-7	GMM	Within treatment only	4	Initially moderate-severe symptoms and fast improvement (21.8%); initially moderate-severe symptoms and gradual improvement (11.9%)		Initially mild symptoms and minimal improvement (41.6%); initially moderate-severe symptoms and no change (24.8%)
Joesch et al. (2013)	482	At baseline, weekly during treatment (12 weeks) and weekly follow-up up to 56 weeks from the baseline	OASIS	GBTM	Within treatment and follow-up	4	Fast improvers with lowest anxiety scores at baseline (27.6%); moderate improvers (40.1%)		Slow improvers with plateau at the end (28.7%); not-improvers with the highest baseline score (3.7%)
Lukaschek et al. (2019)	176	At baseline and 10 periodic within treatment assessments	OASIS	GMM	Within treatment only	3	Improvers, initially high symptoms, constant improvement (33.0%); fast improvement at the beginning and then plateau (50.6%)		Not-improvers (16.5%), an unstable course
Jankowski et al. (2019)	118	No regular frequency reported (at multiple points during and at the end of treatment)	TOP anxiety scale	LGM	Within treatment only	2	Class 1 (21.9%) – higher anxiety levels/reliable improvement		Class 2 (78.1%) lower anxiety levels/no improvement
Altmann et al. (2020)	357	At baseline, at sessions 8/15 and post-treatment	LSAS	LCA	Within treatment only	3	Improvers with severe initial impairment (16%)	Improvers with moderate initial impairment (63%)	Low improvers with severe initial impairment (21%)
H. Kim et al. (2023)	334	At baseline, midway through treatment (week 4) and post-treatment (week 8)	YBOCS	GMM	Within treatment only	3	22.5% dramatic improvement class; 52.1% moderate improvement class		25.4% little to no improvement class
Lutz et al. (2014)	326	At every session	PDSS-SR	GMM	First five sessions only	4	Early response, high symptoms (20.2%)		High symptoms, no change (17.2%); medium symptoms, slow change (58.0%); medium symptoms, early deterioration with further recovery to initial levels (4.6%)
Mean	2522					3.3	40% ^a	1% ^a	59% ^a
Median	334					3.0			

(Continued)

Table 2. Continued.

Study	Sample size	Measurement frequency	Measure used	Modelling method	Period modelled	No. of classes	Responders	Improvers	Non-improvers
<i>Study modelling depression and anxiety in a single model</i>									
Hull et al. (2020)	10,718	At baseline and every three weeks	PHQ-9 and GAD-7	LGM	Within treatment only	6	Acute recovery (7.0%); recovery (23.7%). 19% if averaged for two conditions ^b	Chronic (22.6%); anxiety improvement (16.9%). 32% if averaged for two conditions	Depression improvement (20.0%); elevated chronic (9.8%). 50% if averaged for two conditions
Total for all studies						3.2	36% ^a	9% ^a	55% ^a
Mean	2593								
Median	339					3.0			

OASIS: Overall Anxiety Severity and Impairment Scale; BAI: Beck Anxiety Inventory; BDI: Beck Depression Inventory; BDI-II: Beck Depression Inventory II; GAD-7: Generalized Anxiety Disorder Scale; HDRS: Hamilton Depression Rating Scale; LSAS: Liebowitz Social Anxiety Scale; PDSS-SR: Panic Disorder Severity Scale, self-report form; PHQ-9: Patient Health Questionnaire; SCL-90: Depression Scale of the Symptom Check List90; TOP: Treatment Outcome Package; YBOCS: Yale-Brown Obsessive-Compulsive Scale.

^aAverage weighted by sample size, calculated only for studies included in meta-regression analysis.

^bClassification based on the best improved condition is used in this table, while the average proportion for two conditions is used in meta-analysis. In grey highlight, there are three studies not included in meta-analysis.

($b = -6.8$, $p = 0.003$) were significant predictors, but none of them reached significance in adjusted models.

Research question #2. The initial severity of the condition modelled

For both depression and anxiety, non-improver classes were not always the classes with the highest baseline severity of the condition modelled.

Among the nine depression models with non-improving classes, six included only one non-improving class, and in half of these (Kelley et al., 2018; McDevitt-Petrovic et al., 2020; Siddique et al., 2012), this class had the highest baseline severity. Three studies (Nordberg et al., 2014; Saunders et al., 2019; Skelton et al., 2023) identified two non-improving classes – one characterised by the highest severity and the other by the lowest severity levels.

For anxiety, five models identified one non-improving class, and only in one of them (McDevitt-Petrovic et al., 2020) was it a class with the highest baseline severity. Among the other four studies: one reported two non-improving classes with the highest baseline levels (Joesch et al., 2013), two studies (Saunders et al., 2019; Skelton et al., 2023) reported mixed results (one class had the highest baseline severity and one had either the lowest or average severity) and one study (Lutz et al., 2014) reported three non-improving classes without clear association with baseline severity.

When two conditions were modelled together (Hull et al., 2020), one of two non-improver classes had the highest baseline severity of the symptoms, while the other reported average depression and the lowest anxiety severity.

Table 3 shows that the initial symptom severity level of the condition modelled was one of the most frequently considered potential predictors of class membership, and it received support in over half of the studies (8/11), more often for depression models. Out of those studies, higher initial severity was associated with a less favourable trajectory in about half of the models.

Research question #2. Anxiety and depression mutual class prediction

Three study groups were identified:

1. studies modelling one condition (either anxiety or depression) and measuring the other one at baseline ($k = 13$, 72% of all the studies included);
2. studies modelling both anxiety and depression for the same sample but within separate models ($k = 4$, 22%) (Jankowski et al., 2019; McDevitt-Petrovic et al., 2020; Saunders et al., 2019; Skelton et al., 2023);

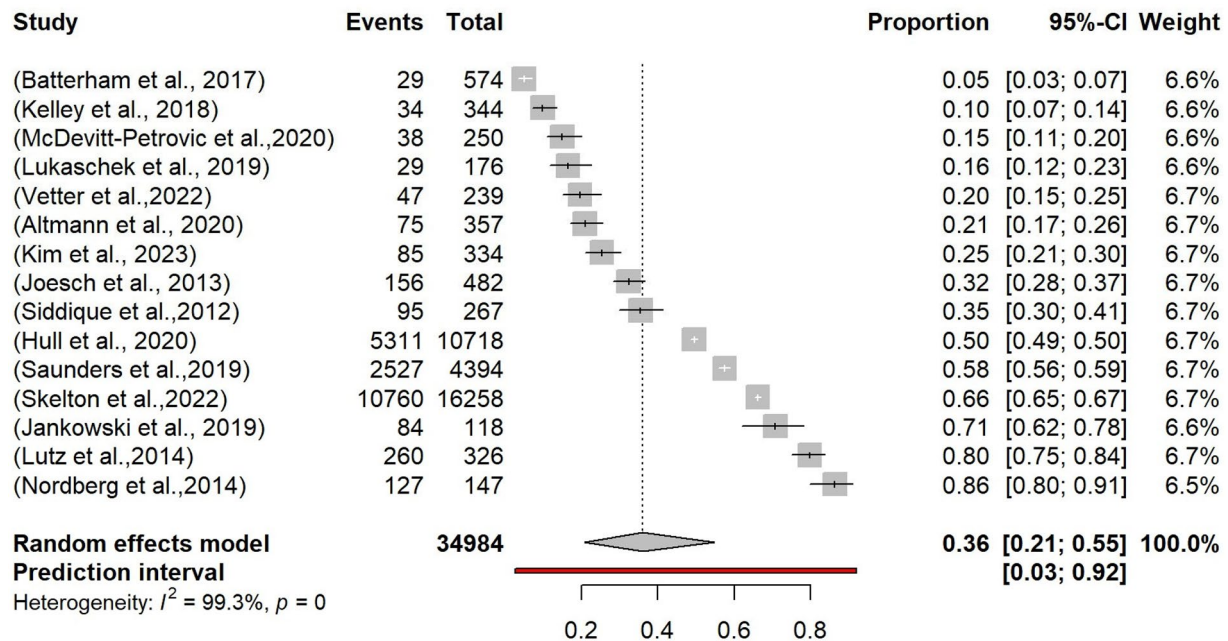


Figure 2. Forest plot of non-improver class proportion.

- studies that analysed anxiety and depression symptom changes within a single model ($k = 1$, 6%) (Hull et al., 2020).

To examine whether initial anxiety and depression symptom severity predicted class allocation for the other condition, we compared the results of 18 models from 15 studies. The third group above was not included due to different methodology; in one study modelling both conditions (Jankowski et al., 2019), no predictor analysis was conducted; in one study (Smolenski et al., 2017), a one-class solution was reported.

Initial anxiety predicted depression class membership in six of 10 models of depression. In one more model (Nordberg et al., 2014), it was a predictor only in univariate analysis and did not reach significance when adjusted for other predictors. Out of those six models, in most ($k = 4$, Cuijpers et al., 2005; Saunders et al., 2019; Skelton et al., 2023; Stulz et al., 2010) higher anxiety predicted allocation to a class with the highest severity of baseline depression; only in one model did it predict non-response vs. response regarding depression symptoms (Vetter et al., 2022); and in one more model (Siddique et al., 2012), it predicted membership in a non-improving class with the highest baseline depression.

Baseline depressive symptoms predicted anxiety class membership in five of eight models of anxiety. In two more models (Joesch et al., 2013; Lutz et al., 2014), they were predictors in univariate analysis only. Out of those five models, in two (Lukaschek et al.,

2019; Saunders et al., 2019), higher initial depression predicted non-response regarding anxiety symptoms; in one model (Altmann et al., 2020), it predicted membership in classes with higher baseline anxiety; in two models, it predicted membership in a non-improving class with the highest baseline anxiety (McDevitt-Petrovic et al., 2020; Skelton et al., 2023).

Table 4 reveals the details of this analysis.

Research question #3. Comparison of individual trajectories across the two conditions

Five studies addressed whether people tend to follow similar trajectory patterns for anxiety and depression during psychological interventions: four (Jankowski et al., 2019; McDevitt-Petrovic et al., 2020; Saunders et al., 2019; Skelton et al., 2023) modelled the symptoms separately, and one (Hull et al., 2020) used a single model.

In three studies (Jankowski et al., 2019; McDevitt-Petrovic et al., 2020; Skelton et al., 2023), the same number of trajectory classes was reported for depression and anxiety, and the class nature was described identically. In one study (Saunders et al., 2019), the authors reported four classes for depression and five classes for anxiety; while the nature was almost the same, some classes differ in terms of initial symptom severity of the symptoms. The shapes of trajectories across the two conditions were visually identical in three studies (Jankowski et al., 2019; Saunders et al., 2019; Skelton et al., 2023); the authors of the fourth one (McDevitt-Petrovic et al., 2020) reported no plot, but the slope coefficients were almost identical across

Table 3. Most frequently considered predictors for trajectory class membership (ordered according to the number of studies that considered and supported a predictor).

Predictor	Studies considered a predictor		Studies supported a predictor			Models considered/supported		Comments
	<i>K</i>	Average sample size	<i>k</i>	% in studies considered	Average sample size	Depression	Anxiety	
Gender/Sex	13	2689	3	23%	9051	7/0	8/2	Being female/woman was associated with membership of classes with higher baseline severity (Hull et al., 2020; Lukaschek et al., 2019; Skelton et al., 2023)
Age	13	2683	2	15%	406	7/2	8/0	Younger age was a predictor of non-response (Batterham et al., 2017; Vetter et al., 2022)
Initial severity of the condition modelled	11	703	8	73%	841	8/7	4/2	In four models of depression (Batterham et al., 2017; Saunders et al., 2019; Siddique et al., 2012; Vetter et al., 2022) initial severity level was a predictor of non-improver class membership; in three models it was a predictor of intercept only (Cuijpers et al., 2005; Nordberg et al., 2014; Stulz et al., 2010). For anxiety models: Saunders et al. (2019) reported baseline severity to predict non-response and Lukaschek et al. (2019) reported it to distinguish classes' intercepts, not differences in response
Education	7	1840	2	29%	5538	2/0	4/1	High education predict better improvement (Hull et al., 2020) and/or low baseline severity (Altmann et al., 2020; Hull et al., 2020)
Relationship status	7	394	2	29%	298	4/1	3/1	Controversial results: no close relationship predicted non-improver membership vs. responding one, both with high baseline anxiety (Altmann et al., 2020); married status was a predictor of non-improver vs. improver class, both classes with high baseline depression (Vetter et al., 2022)
Treatment type	6	2102	5	83%	2422	4/3	1/1	Mixed results: CBT subjects responded worse when compared to patients treated with medication (Kelley et al., 2018; Siddique et al., 2012) but responded better when compared to ones treated as usual if they have higher baseline severity (Cuijpers et al., 2005). CBT treatment (vs. PDT) or therapist's CBT orientation predicted higher chances of response (Altmann et al., 2020; Hull et al., 2020)
Employment	6	3013	4	67%	4307	5/2	3/3	Not-employed/not working status is a predictor of non-improver class membership for both depression (Skelton et al., 2023; Vetter et al., 2022) and anxiety (Joesch et al., 2013; McDevitt-Petrovic et al., 2020; Skelton et al., 2023)
Medication	6	3663	3	50%	5671	5/2	4/2	In depression models medication prescribed or used before decreased odds of fast improvement (Skelton et al., 2023) or of having lower baseline severity (Skelton et al., 2023; Stulz et al., 2010). In anxiety models, it was associated with non-response (McDevitt-Petrovic et al., 2020; Skelton et al., 2023)
Ethnicity	6	3684	2	33%	8301	5/2	3/1	Mixed results. In depression models: Asian, Black and mixed background (vs. White) was associated with lower odds to improve faster or with lower baseline class (Skelton et al., 2023); Hispanics were more likely to respond earlier when compared to non-Hispanic (Kelley et al., 2018). For anxiety: Black background was associated with lower chances to gradually improve vs. not improve (Skelton et al., 2023)
Social functioning	4	5370	3	75%	6993	3/2	3/3	Higher impairment decreased odds of being in faster improving classes (Saunders et al., 2019; Skelton et al., 2023) or classes with lower baseline severity (Lutz et al., 2014; Saunders et al., 2019; Skelton et al., 2023)
Onset of condition modelled	4	358	1	25%	250	3/1	2/0	An onset of >5 years increased the probability of being in the non-improvers class vs. responding class (McDevitt-Petrovic et al., 2020)
Suicidal ideation/suicidality	3	320	2	67%	193	3/2	0/0	Suicidality was associated with non-response (Nordberg et al., 2014; Vetter et al., 2022)
History of psychotherapy	3	363	0	0%	–	2/0	2/0	No statistically significant results

CBT: cognitive behavioural therapy; PDT: psychodynamic psychotherapy.

Model split into depression and anxiety does not include the study modelling both conditions together (Hull et al., 2020). A predictor was coded as significant in case it was reported to be in multivariate analysis (if it was significant in univariate analysis only, it was coded as not significant for this table).

Table 4. Results of analysis of predictors for trajectory class allocation (ordered according to the results reported: non-response prediction comes first).

Study	Sample size	No. of classes	Depression and anxiety mutual class prediction		Other significant predictors	Method used
			Results reported	Details of results		
<i>Studies modelling depression</i>						
Vetter et al. (2022)	239	3	Predicts non-response only	Anxiety is a predictor of non-improving class allocation (with lower baseline depression) vs. improving (with highest baseline depression)	Most relevant predictors for non-response were insomnia symptoms, younger age, depressed mood, being unemployed, suicidal ideation, somatic symptoms of depressive disorder, being married	The binary classification model (eXtreme Gradient Boosting; XGBoost). Only two classes (improving vs. non-improving) were compared
Siddique et al. (2012)	267	2	Predicts both level and non-response ^a	Baseline anxiety is associated with higher odds to be allocated to severe depression non-improving class (vs. moderate depression improving class)	Baseline depression scores (high-symptom non-improvers vs. moderate-symptom improvers); type of treatment	Logistic regression and a classification tree, conditional GMM for treatment type
Saunders et al. (2019)	4394	4	Predicts level only	Higher initial anxiety score is a statistically significant predictor of depression non-responding class with highest baseline depression vs. all other classes (two responding and one non-improving with lowest but still high baseline depression)	Non-responding trajectories were associated with higher baseline severity, worse social functioning and higher incidence of phobic anxiety	Multinomial logistic regression
Skelton et al. (2023)	16,258	4	Predicts level only	Initial anxiety score is a predictor of a non-improving class with mild initial depression membership vs. a non-improving class with high initial depression. But it is not for two responding classes with high depression vs. a non-improving class with high depression	Non-improving class membership was associated with disability and unemployed status. Psychotropic medication prescribed decreased the odds of belonging to fast improvers and the class with mild baseline depression. Other predictors: baseline impairment, ethnicity, comorbidity	Multinomial logistic regression
Stulz et al. (2010)	504	3	Predicts level only	Higher baseline anxiety scores predict a higher probability of belonging to the severe depression subgroup (as opposed to the moderate depression subgroup). Both are responders	History of antidepressant medication, IDS-SR score, and FAD score	Multinomial logistic regression, conditional GMM for treatment type
Cuijpers et al. (2005)	425	4	Predicts level only	Anxiety disorder is a significant predictor for odds of allocation to two classes with higher baseline depression vs. a class with lowest baseline depression	A severe MDD and a dysthymic disorder (prediction of the allocation to a class with the highest baseline depression); type of treatment	Conditional GMM
Nordberg et al. (2014)	147	3	Not enough information	The panic subscale is a significant predictor in univariate analysis (no details reported) but not in multivariate	The depression subscale was a predictor of low-symptom non-responders vs. high-symptom responders. Increased social conflict, poorer sexual functioning, increased suicidality and lower hostility were associated with increased likelihood of belonging to the high-symptom non-responder group vs. high-symptom responding group	Logistic regression
McDevitt-Petrovic et al. (2020)	250	2	Not significant	–	An onset of >5 years and receiving concurrent treatment significantly increased the probability of being in the non-improvers class	Chi-square tests and binary logistic regression
Batterham et al. (2017)	574	2	Not significant	–	Factors significantly associated with membership of the stable/deteriorating depression class were higher baseline depression scores, younger age and limited comfort with using the internet	Binary logistic regression
Kelley et al. (2018)	344	4	Not significant	–	Ethnicity, type of treatment (CBT vs. medication), genetic factors	Ordinal logistic regression and ANOVA
Mean	3681 329	– –				

(Continued)

Table 4. Continued.

Study	Sample size	No. of classes	Depression and anxiety mutual class prediction		Other significant predictors	Method used
			Results reported	Details of results		
<i>Studies modelling anxiety</i>						
Lukaschek et al. (2019)	176	3	Predicts non-response	Statistically significant differences between classes in %% of MDD (the highest % is reported for a non-improving class while it is not the highest in terms of baseline anxiety)	Sex, anxiety baseline score, rural/urban residence	ANOVA
Saunders et al. (2019)	4394	5	Predicts non-response	A higher initial depression score is a statistically significant predictor of anxiety non-responding class (with the highest initial anxiety) vs. 3 responding classes but not vs. another non-improving class (with lower but still high initial anxiety)	Non-responding trajectories were associated with higher baseline severity, worse social functioning and higher social phobia and specific phobia scores	Multinomial logistic regression
McDevitt-Petrovic et al. (2020)	250	2	Predicts both level and non-response ^a	Comorbidity significantly increases the probability of being in the non-improvers class (with higher baseline anxiety)	Unemployment and using medication significantly increased the probability of being in the non-improvers class	Chi-square tests and binary logistic regression
Altmann et al. (2020)	357	3	Predicts level only	Depression is a significant predictor of the allocation to an improving class with moderate baseline anxiety vs. two classes with high baseline anxiety (one responding and one non-improving)	Education, relationship status, anxiety in relationship, therapy type, therapist in training, alliance, interpersonal problems	Multinomial logistic regression
Skelton et al. (2023)	16,258	4	Mixed results	Initial depression score is a predictor of membership in two classes (responding from high anxiety and non-improving from mild anxiety) vs. non-responding class with high anxiety. No difference is reported for responding from high anxiety class vs. non-improving from high anxiety class	No-change class was associated with psychotropic medication prescription, disability and unemployment status. Fast improvers and class with lower baseline anxiety were associated with lower baseline impairment. Being a male was associated with the class with lower baseline anxiety. Other predictors: ethnicity, comorbidity	Multinomial logistic regression
Joesch et al. (2013)	482	4	Predicts both level and non-response ^a but only when unadjusted for other predictors	Baseline depression is a significant predictor of membership in two non-improving groups with the highest baseline anxiety (and not in a responding group with lower anxiety) vs. a responder group with the lowest anxiety membership. In multivariate analysis it failed to reach statistical significance	The number of anxiety disorders, the extent of social support and not working	Univariate and multivariate logistic regression
Lutz et al. (2014)	326	4	Mixed results, only when unadjusted for other predictors	The baseline depression score is significantly lower for two classes with lower baseline anxiety levels compared to two classes with high anxiety in univariate analysis, but not statistically significant in multivariate analysis	Impairment in social functioning and separation anxiety	ANOVA, multinomial logistic regression
Kim et al. (2023)	334	3	Not significant	–	YBOCS avoidance scores and transdiagnostic internalizing factor scores	Multinomial logistic regression
Mean	4287 381	– –				

ANOVA: analysis of variance; MDD: major depression disorder; YBOCS: Yale-Brown Obsessive-Compulsive Scale.

^aModels where classes with higher baseline severity follow less favourable trajectories and thus it is impossible to distinguish if a comorbid condition predicts non-response (less favourable trajectory class) or the baseline severity level of the condition modelled.

the two conditions for improving classes and close for non-improving classes.

The analysis of overlapping samples (proportions of participants following trajectories of the same nature for both depression and anxiety) was conducted in two studies (Saunders et al., 2019; Skelton et al., 2023):

- most depression non-improvers followed the non-improver class(-es) on anxiety as well (89–90%); but for anxiety non-improvers, the spread over depression classes was more diversified: only 48–61% of them followed the non-improving depression class, while the rest demonstrated different rates of improvement in depressive symptoms;
- the overlap between the opposite classes, such as anxiety non-improvers and depression rapid improvers, was the lowest at 1–2%.

One study (Jankowski et al., 2019) reported nothing about the overlap; the other one (McDevitt-Petrovic et al., 2020) reported correlation ratios between anxiety and depressive symptoms on session-by-session basis (all the ratios were above 0.7).

In one single-model study (Hull et al., 2020), participants were allocated to trajectory classes based on joint patterns of the two conditions:

- participants of four classes (63% of the sample) followed similar depression and anxiety starting from the same baseline levels: i.e. those with severe depressive symptoms also had high anxiety scores;
- participants of two classes demonstrated more nuanced comorbidity at baseline with either depression or anxiety prevailing and a better improvement of the more severe condition.

Research question #4. Other predictors

Fifteen studies informed results of other predictor analysis as revealed by Table 3.

Age and gender/sex were considered as potential predictors of class membership in most studies but reached statistical significance in only 15% and 23% of cases, respectively. More consistent predictors of non-response included unemployment and higher suicidality, while social functioning impairment was associated with lower response or higher initial severity of the condition modelled. Treatment type was a significant predictor in five of six studies: CBT was associated with less favourable trajectories than medication but more favourable than TAU and PDT. In half of the

studies considering medication (prescribed or used before), it increased the chances of following either a non-improving trajectory or trajectory with higher baseline severity.

Predictor analyses for depression and anxiety were largely consistent, with two exceptions: (i) unemployment more strongly predicted non-response in anxiety models and (ii) relationship status showed conflicting effects in anxiety and depression models.

Discussion

We identified 22 models of depressive and/or anxiety symptom changes in the adult population of 18 studies and reclassified trajectory classes into responders, improvers and non-improvers. The number of classes extracted was associated with the sample size. Most classes were responders and non-improvers. For these two classes, we conducted meta-regression of the class proportion, which revealed no associations with any of the study characteristics we analysed. In most studies, initial levels of depressive and anxiety symptoms predicted class membership both in terms of the same condition and the other one (i.e. initial depression level predicting anxiety class membership). However, higher initial severity of the symptoms was not always associated with non-improving vs. responding class membership. Most adults tended to follow the same recovery trajectories regarding both anxiety and depression. Among other predictors, unemployment, suicidality, social functioning impairment, treatment type and medication use appeared to be the most likely risk factors for membership in a non-improving class.

Number and nature of trajectories and their association with study characteristics

The number of trajectory classes extracted was in the range from two to six, with an average of 3.25, which is consistent with the existing reviews (Cole, 2021; Dewar et al., 2020). Its association with the sample size is also echoed by findings reported by Cole (2021) based on 45 studies on CMD trajectories. The association highlights the need for greater statistical power to extract more classes, as simulation studies show that identifying more true classes requires larger samples (S.-Y. Kim, 2012). Meanwhile, there is no clear rule for extracting a specific number of classes, and the extraction depends on numerous factors (Ram & Grimm, 2009). Therefore, we used no sample size cut-offs for the studies reviewed but provided, where applicable, references to the average sample size.

Among other study characteristics, the sample type was associated with the number of classes after controlling for the sample size. Clinical sample in this review was defined as a sample in which participants were initially screened as having symptom severity scores above a cut-off on a relevant measure, or formally diagnosed with the modelled condition. Studies with only clinical populations demonstrated a higher mean number of classes than studies with mixed populations. Further research is required to explore the differences of trajectories in clinical and subclinical populations, as our review is limited by the number of eligible studies.

We applied the RCI and clinically significant change approach to classify trajectory classes extracted into responders, improvers and non-improvers. Meta-analysis of the pooled proportion of responding and not improving classes reported high levels of I^2 values, which is common for meta-analyses of prevalence (Migliavaca et al., 2022). In line with the recommendations (Borges Migliavaca et al., 2020), we also reported extremely wide prediction intervals that confirmed the heterogeneity of the pooled estimates.

Meta-regression of depression models observed that the responding class proportion was higher in samples with a lower proportion of females and when the follow-up period was not included. The adjusted meta-regression model did not support the associations, possibly due to the limited number of studies. While the association with follow-up inclusion can be explained by the diminishing effect of interventions, sex effect on intervention effectiveness is not supported by the existing research: while women tend to report higher severity of depression (Buckman et al., 2021) meta-analytical studies failed to find significant differences in psychotherapy effects for women only and mixed adult groups (Cuijpers et al., 2018).

Initial levels of depression and anxiety

We found that the baseline severity of the condition modelled was not always associated with non-improving class membership. Only about half of depression non-improving classes and about 1/3 of anxiety ones were classes with the highest baseline severity. The baseline symptom severity was regularly identified as a predictor for that symptom class membership; however, higher severity was associated with being a member of a non-improving class only in half of depression models and in just one of four models of anxiety. Thus, neither line of analysis supported the initial severity predicting non-improvement. Cole (2021) and Dewar et al. (2020) also reported mixed results.

Our analysis revealed strong support for anxiety and depression to predict class membership of each other. Studies with larger sample sizes reported this association more frequently, which may be explained by the insufficient statistical power of smaller sample studies (de Jong et al., 2019). Meanwhile, a higher comorbid symptom predicted a non-improvement only in one depression model and two anxiety models; in the remaining models, it predicted allocation to a class with either a higher baseline severity of the symptom modelled or both a higher severity and non-improvement.

These findings illustrate the more complex effects of baseline severity of both a condition modelled and a comorbid one on the response curve. The interaction of the two effects is also possible. Baseline depression severity (compared to anxiety) finds more support to serve as a predictor of non-response both when depression was modelled and when it was considered as a predictor for anxiety trajectory allocation. This is also echoed by the existing research showing that higher baseline severity of depression is associated with poorer eventual prognosis in treatment (Buckman et al., 2021). Further analysis with a larger sample of studies is required to verify these associations, and the mechanism of the conditions' interplay needs further investigation.

Comparison of trajectories for depression and anxiety

Five studies informed the analysis of the differences in shapes of trajectories across anxiety and depression modelled for the same sample. When the conditions were modelled separately, a considerable overlap between the classes of similar change patterns was reported, i.e. non-improving participants in terms of one condition tend to be allocated to a non-improver class in terms of the other condition. When modelled in a single model, the majority of participants were allocated to classes with the same pattern of change of both anxiety and depression. This tendency is supported by existing research that revealed medium to high secondary effects of psychological interventions delivered for depression on anxiety symptoms (Weitz et al., 2018) and for GAD on depressive symptoms (Cuijpers et al., 2014).

Other predictors of class membership

Unemployed status and higher suicidal ideation were the most frequent predictors of following non-improvement

trajectories; impairment in social functioning and medication predicted either non-response or higher baseline severity of a condition modelled. This echoes the previously reported systematic reviews (Cole, 2021; Dewar et al., 2020) and other research in the area. Thus, meta-analytical research reported a history of antidepressant treatment predicting worse outcomes of depression treatment (Buckman et al., 2021); the same study linked unemployment to higher baseline depression severity, which, in turn, predicted poorer treatment outcomes.

Six studies considered the type of treatment as a potential predictor of trajectory class allocation. They reported CBT being more effective compared to other non-medication treatments but underperforming medication treatments. Meanwhile, the number of these studies was limited, and this association is not supported by the results reported by Cole (2021). Moreover, no superior effectiveness of CBT was revealed by meta-analytical studies either for GAD (Cuijpers et al., 2014) or for depression (Cuijpers et al., 2023). We also failed to find any effects of CBT-based interventions on the proportion of non-improving or responding classes in the meta-regression, although this may be explained methodologically: when the type of intervention – including CBT-based – was analysed within conditional growth mixture model (three studies used this method), we had to code the intervention as mixed rather than CBT-based.

Unlike previous findings for PTSD (Dewar et al., 2020), our review did not support the conclusion that younger age predicts membership in the improver versus non-improver class. Although age was the most frequently examined risk factor, it was only found to be associated in two of 13 studies included, with these indicating that younger age was associated with non-response. Our results align with the recent meta-analysis by Cuijpers et al. (2020), which analysed 366 randomised clinical trials on psychological interventions for depression and found no significant differences in treatment effects between middle-aged patients and older adults. Similar findings were previously reported for GAD (Cuijpers et al., 2014).

When comparing predictors across depression and anxiety, the results were largely consistent.

Reporting quality

We assessed the reporting quality of the studies using the GRoLTS checklist (van de Schoot et al., 2017). The low average total score might be partly due to

the checklist's recent introduction. However, we found no association between reporting quality and the year of publication. No studies were excluded based on a low score, but the GRoLTS score was strongly associated with sample size, suggesting that studies with larger samples may be more reliable.

To improve future research in this field, we recommend adopting the GRoLTS checklist. Its use will not only enhance replicability but also facilitate meta-analytical research, which is currently hindered by the exploratory nature of studies, methodological diversity, and inconsistencies in reported trajectory classes.

Limitations

There are some limitations to acknowledge. First, the inclusion criteria and our focus on the interplay of anxiety and depressive symptoms (namely, the requirement for both conditions to be measured) limited the number of studies included. Due to the diversity of methods used and studies' characteristics, the application of meta-analysis was limited, and a narrative synthesis was the key synthesis approach. Moreover, only one study modelled both conditions in the same model, making it impossible to compare results between separate- and single-model approaches. Second, while using the RCI and clinically significant change approach to classify trajectory classes reduced subjectivity in labelling classes, some bias may still be present. Clear and widely accepted guidelines were not available for all the measures, making our classification dependent on the samples we used to calculate the RCI and clinical cut-offs when they were unavailable. Third, data extraction was performed by a single researcher with an independent rater involved in a partial quality assessment only. Finally, only studies in English and from high-income countries were included, limiting the generalisability of the results.

Conclusion

To the best of our knowledge, this review is the first to consider the interplay between anxiety and depression symptom changes during psychological interventions. Our findings suggest that the initial severity of both the condition modelled and the comorbid one may predict the trajectory class membership; however, higher initial severity does not always predict following a non-improving trajectory. Unemployed status, suicidality, social functioning impairment, treatment type and medication use received the strongest

support among other predictors of non-improvement. Adults tend to follow similar trajectories in terms of depressive and anxiety symptoms.

The results can inform practitioners around estimating potential treatment response for adults with comorbid depression and anxiety. Beyond identifying factors associated with non-response, our findings indicate the need to look beyond initial symptom severity and highlight the mutual influence of the two conditions. Applying methods like GMM to treatment response data can reveal expected patterns of change for clients based on their presenting characteristics. This could support client monitoring and progress feedback, for example, identifying individuals who are “off-track”, which was shown to improve outcomes (Shimokawa et al., 2010).

Acknowledgements

Our appreciation goes to Anna Kennedy for her role as the second rater. Her independent assessments were invaluable in ensuring the consistency and accuracy of the studies’ reporting quality evaluation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Anna Kamardina  <http://orcid.org/0000-0002-1350-4510>

Rob Saunders  <http://orcid.org/0000-0002-7077-8729>

Emma Broglia  <http://orcid.org/0000-0003-2137-8187>

References

- Altmann, U., Gawlytta, R., Hoyer, J., Leichsenring, F., Leibing, E., Beutel, M., Willutzki, U., Herpertz, S., & Strauss, B. (2020). Typical symptom change patterns and their predictors in patients with social anxiety disorder: A latent class analysis. *Journal of Anxiety Disorders*, 71, 102200. <https://doi.org/10.1016/j.janxdis.2020.102200>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. American Psychiatric Association.
- Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health*, 22(4), 153–160. <https://doi.org/10.1136/ebmental-2019-300117>
- 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>
- Borges Miglavaca, C., Stein, C., Colpani, V., Barker, T. H., Munn, Z., & Falavigna, M. Prevalence Estimates Reviews – Systematic Review Methodology Group (PERSys). (2020). How are systematic reviews of prevalence conducted? A methodological study. *BMC Medical Research Methodology*, 20(1), 96. <https://doi.org/10.1186/s12874-020-00975-3>
- Buckman, J. E. J., Saunders, R., Cohen, Z. D., Barnett, P., Clarke, K., Ambler, G., DeRubeis, R. J., Gilbody, S., Hollon, S. D., Kendrick, T., Watkins, E., Wiles, N., Kessler, D., Richards, D., Sharp, D., Brabyn, S., Littlewood, E., Salisbury, C., White, I. R., ... Pilling, S. (2021). The contribution of depressive ‘disorder characteristics’ to determinations of prognosis for adults with depression: An individual patient data meta-analysis. *Psychological Medicine*, 51(7), 1068–1081. <https://doi.org/10.1017/S0033291721001367>
- Cole, C. L. (2021). *Do symptom change trajectories differ between cognitive behavioural therapy (CBT) and counselling? An application of growth curve modelling using a large IAPT dataset*. University College London (UCL).
- Core Team, R. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Cuijpers, P., Karyotaki, E., Eckshtain, D., Ng, M. Y., Corteselli, K. A., Noma, H., Quero, S., & Weisz, J. R. (2020). Psychotherapy for depression across different age groups: A systematic review and meta-analysis. *JAMA Psychiatry*, 77(7), 694–702. <https://doi.org/10.1001/jama-psychiatry.2020.0164>
- Cuijpers, P., Karyotaki, E., Reijnders, M., & Huibers, M. J. H. (2018). Who benefits from psychotherapies for adult depression? A meta-analytic update of the evidence. *Cognitive Behaviour Therapy*, 47(2), 91–106. <https://doi.org/10.1080/16506073.2017.1420098>
- Cuijpers, P., Miguel, C., Ciharova, M., Harrer, M., Basic, D., Cristea, I. A., de Ponti, N., Driessen, E., Hamblen, J., Larsen, S. E., Matbouriahi, M., Papola, D., Pauley, D., Plessen, C. Y., Pfund, R. A., Setkowski, K., Schnurr, P. P., van Ballegooijen, W., Wang, Y., ... Karyotaki, E. (2024). Absolute and relative outcomes of psychotherapies for eight mental disorders: A systematic review and meta-analysis. *World Psychiatry*, 23(2), 267–275. <https://doi.org/10.1002/wps.21203>
- Cuijpers, P., Miguel, C., Harrer, M., Plessen, C. Y., Ciharova, M., Papola, D., Ebert, D., & Karyotaki, E. (2023). Psychological treatment of depression: A systematic overview of a ‘meta-analytic research domain’. *Journal of*

- Affective Disorders*, 335, 141–151. <https://doi.org/10.1016/j.jad.2023.05.011>
- Cuijpers, P., Sijbrandij, M., Koole, S., Huibers, M., Berking, M., & Andersson, G. (2014). Psychological treatment of generalized anxiety disorder: A meta-analysis. *Clinical Psychology Review*, 34(2), 130–140. <https://doi.org/10.1016/j.cpr.2014.01.002>
- Cuijpers, P., Van Lier, P. A. C., 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>
- de Jong, V. M. T., Eijkemans, M. J. C., van Calster, B., Timmerman, D., Moons, K. G. M., Steyerberg, E. W., & van Smeden, M. (2019). Sample size considerations and predictive performance of multinomial logistic prediction models. *Statistics in Medicine*, 38(9), 1601–1619. <https://doi.org/10.1002/sim.8063>
- Delgadillo, J., & Gonzalez Salas Duhne, P. (2020). Targeted prescription of cognitive-behavioral therapy versus person-centered counseling for depression using a machine learning approach. *Journal of Consulting and Clinical Psychology*, 88(1), 14–24. <https://doi.org/10.1037/ccp0000476>
- 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. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, 65(2), 71–86. <https://doi.org/10.1177/0706743719875602>
- 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>
- 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>
- Jankowski, P. J., Sandage, S. J., Bell, C. A., Rupert, D., Bronstein, M., & Stavros, G. S. (2019). Latent trajectories of change for clients at a psychodynamic training clinic. *Journal of Clinical Psychology*, 75(7), 1147–1168. <https://doi.org/10.1002/jclp.22769>
- 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>
- Kelley, M. E., Dunlop, B. W., Nemeroff, C. B., Lori, A., Carrillo-Roa, T., Binder, E. B., Kutner, M. H., Rivera, V. A., Craighead, W. E., & Mayberg, H. S. (2018). Response rate profiles for major depressive disorder: Characterizing early response and longitudinal nonresponse. *Depression and Anxiety*, 35(10), 992–1000. <https://doi.org/10.1002/da.22832>
- Kessler, R. C., Sampson, N. A., Berglund, P., Gruber, M. J., Al-Hamzawi, A., Andrade, L., Bunting, B., Demyttenaere, K., Florescu, S., de Girolamo, G., Gureje, O., He, Y., Hu, C., Huang, Y., Karam, E., Kovess-Masfety, V., Lee, S., Levinson, D., Medina Mora, M. E., ... Wilcox, M. A. (2015). Anxious and non-anxious major depressive disorder in the World Health Organization World Mental Health Surveys. *Epidemiology and Psychiatric Sciences*, 24(3), 210–226. <https://doi.org/10.1017/S2045796015000189>
- Kim, H., Wheaton, M. G., Foa, E. B., & Simpson, H. B. (2023). Identifying trajectories of symptom change in adults with obsessive compulsive disorder receiving exposure and response prevention therapy. *Journal of Anxiety Disorders*, 96, 102711. <https://doi.org/10.1016/j.janxdis.2023.102711>
- Kim, S.-Y. (2012). Sample size requirements in single- and multiphase growth mixture models: A Monte Carlo simulation study. *Structural Equation Modeling*, 19(3), 457–476. <https://doi.org/10.1080/10705511.2012.687672>
- Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, 22(17), 2693–2710. <https://doi.org/10.1002/sim.1482>
- Lukaschek, K., Hiller, T. S., Schumacher, U., Teismann, T., Breitbart, J., Brettschneider, C., König, H.-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(1), 7170. <https://doi.org/10.1038/s41598-019-43487-x>
- 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>
- McCrone, P. R. (Ed.). (2008). *Paying the price: The cost of mental health care in England to 2026*. King's Fund.
- 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. *British Journal of Clinical Psychology*, 59(2), 169–185. <https://doi.org/10.1111/bjc.12237>
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282. <https://doi.org/10.11613/bm.2012.031>
- Migliavaca, C. B., Stein, C., Colpani, V., Barker, T. H., Ziegelmann, P. K., Munn, Z., & Falavigna, M., Prevalence Estimates Reviews—Systematic Review Methodology Group (PERSyst). (2022). Meta-analysis of prevalence: I^2 statistic and how to deal with heterogeneity. *Research Synthesis Methods*, 13(3), 363–367. <https://doi.org/10.1002/jrsm.1547>
- Moller, N. P., Ryan, G., Rollings, J., & Barkham, M. (2019). The 2018 UK NHS Digital annual report on the

- Improving Access to Psychological Therapies programme: A brief commentary. *BMC Psychiatry*, 19(1), 252. <https://doi.org/10.1186/s12888-019-2235-z>
- Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class-latent growth modeling. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 291–322). American Psychological Association.
- Muthén, B. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (ed.), *The SAGE handbook of quantitative methodology for the social sciences* (pp. 346–369). SAGE Publications, Inc. <https://doi.org/10.4135/9781412986311.n19>
- Newman, M. G., Shin, K. E., & Lanza, S. T. (2019). Time-varying moderation of treatment outcomes by illness duration and comorbid depression in generalized anxiety disorder. *Journal of Consulting and Clinical Psychology*, 87(3), 282–293. <https://doi.org/10.1037/ccp0000385>
- 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: Rapid responders. *Journal of Clinical Psychology*, 70(9), 886–903. <https://doi.org/10.1002/jclp.22077>
- Nye, A., Delgadillo, J., & Barkham, M. (2023). Efficacy of personalized psychological interventions: A systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology*, 91(7), 389–397. <https://doi.org/10.1037/ccp0000820>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- 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>
- Ruscio, A. M., Hallion, L. S., Lim, C. C. W., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Andrade, L. H., Borges, G., Bromet, E. J., Bunting, B., Caldas de Almeida, J. M., Demyttenaere, K., Florescu, S., de Girolamo, G., Gureje, O., Haro, J. M., He, Y., Hinkov, H., Hu, C., ... Scott, K. M. (2017). Cross-sectional comparison of the epidemiology of DSM-5 generalized anxiety disorder across the globe. *JAMA Psychiatry*, 74(5), 465–475. <https://doi.org/10.1001/jamapsychiatry.2017.0056>
- 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., Cape, J., Fearon, P., & Pilling, S. (2016). Predicting treatment outcome in psychological treatment services by identifying latent profiles of patients. *Journal of Affective Disorders*, 197, 107–115. <https://doi.org/10.1016/j.jad.2016.03.011>
- Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology*, 78(3), 298–311. <https://doi.org/10.1037/a0019247>
- 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>
- Skelton, M., Carr, E., Buckman, J. E. J., Davies, M. R., Goldsmith, K. A., Hirsch, C. R., Peel, A. J., Rayner, C., Rimes, K. A., Saunders, R., Wingrove, J., Breen, G., & Eley, T. C. (2023). Trajectories of depression and anxiety symptom severity during psychological therapy for common mental health problems. *Psychological Medicine*, 53(13), 6183–6193. <https://doi.org/10.1017/S0033291722003403>
- Smolenski, D. J., Pruitt, L. D., Vuletic, S., Luxton, D. D., & Gahm, G. (2017). Unobserved heterogeneity in response to treatment for depression through videoconference. *Psychiatric Rehabilitation Journal*, 40(3), 303–308. <https://doi.org/10.1037/prj0000273>
- Steel, Z., Marnane, C., Iranpour, C., 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>
- 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 re-analysis. *Journal of Consulting and Clinical Psychology*, 78(3), 409–419. <https://doi.org/10.1037/a0019267>
- Tunvirachaisakul, C., Gould, R. L., Coulson, M. C., Ward, E. V., Reynolds, G., Gathercole, R. L., Grocott, H., Supasitthumrong, T., Tunvirachaisakul, A., Kimona, K., & Howard, R. J. (2018). Predictors of treatment outcome in depression in later life: A systematic review and meta-analysis. *Journal of Affective Disorders*, 227, 164–182. <https://doi.org/10.1016/j.jad.2017.10.008>
- 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>
- Vetter, J. S., Schultebrucks, K., Galatzer-Levy, I., Boeker, H., Brühl, A., Seifritz, E., & Kleim, B. (2022). Predicting non-response to multimodal day clinic treatment in severely impaired depressed patients: A machine learning

- approach. *Scientific Reports*, 12(1), 5455. <https://doi.org/10.1038/s41598-022-09226-5>
- Weitz, E., Kleiboer, A., Van Straten, A., & Cuijpers, P. (2018). The effects of psychotherapy for depression on anxiety symptoms: A meta-analysis. *Psychological Medicine*, 48(13), 2140–2152. <https://doi.org/10.1017/S0033291717003622>
- Wolitzky-Taylor, K. B., Arch, J. J., Rosenfield, D., & Craske, M. G. (2012). Moderators and non-specific predictors of treatment outcome for anxiety disorders: A comparison of cognitive behavioral therapy to acceptance and commitment therapy. *Journal of Consulting and Clinical Psychology*, 80(5), 786–799. <https://doi.org/10.1037/a0029418>