Article

César González-Blanch[1,2,3,4,*]
María Camino-Sánchez[5]
Rosario Bengochea-Seco[4,5]
Cintia Montes-Novoa[5]
David Gil-Sanz[4,5]
Silvia Pérez-Monzón[5]
Victoria López-Sánchez[5]
Paula Díaz-Gómez[6,7]
Amador Priede[2,6,8]

Associations of Baseline Depression and Anxiety Severity With Reliable Improvement and Recovery in Primary Care Psychological Therapy

[1]Mental Health Centre, University Hospital Marqués de Valdecilla, 39008 Santander, Spain
[2]IDIVAL, Valdecilla Biomedical Research Institute, 39011 Santander, Spain
[3]Department of Psychology, International University of La Rioja (UNIR), 26006 Logroño, Spain
[4]Faculty of Health Sciences, European University of the Atlantic, 39011 Santander, Spain
[5]Primary Care Services, Cantabrian Health Service (SCS), 39011 Cantabria, Spain
[6]Mental Health Strategy and Planning Office, Cantabrian Health Service (SCS), 39011 Cantabria, Spain
[7]Department of Nursing, University Hospital Marqués de Valdecilla, 39011 Santander, Spain
[8]Mental Health Centre, Hospital General Sierrallana, 39300 Torrelavega, Spain

Abstract

Background: Brief psychological treatment programmes in primary care are increasingly used to address common mental disorders (CMDs), yet patient outcomes remain heterogeneous. Routine symptom measures may help identify patients who are more or less likely to achieve clinically meaningful change and may support treatment planning within stepped-care systems. This study examined associations between baseline depressive and anxiety symptom severity predicted reliable improvement and reliable recovery following brief psychological treatment in a real-world primary care programme.

Methods: This secondary analysis used routinely collected data from 512 adults treated in a real-world primary care psychological intervention programme. Participants were assessed with the Patient Health Questionnaire-9 (PHQ-9) and Generalized Anxiety Disorder-7 (GAD-7) before and after brief psychological treatment. Baseline demographic and clinical variables, including age, sex, number of sessions, and PHQ-9 and GAD-7 scores, were examined according to reliable improvement and reliable recovery status. Logistic regression analyses examined associations between baseline variables and treatment outcomes.

Results: Reliable improvement was observed in 73.6% of participants, whereas 53.3% achieved reliable recovery. Higher baseline anxiety severity was associated with reliable improvement (odds ratio [OR] = 1.16, 95% confidence interval [CI]: 1.08–1.23), whereas lower depressive severity predicted reliable recovery (OR = 0.92, 95% CI: 0.88–0.96). Fewer treatment sessions were associated with both reliable improvement and reliable recovery. Age and sex were not significantly associated with outcomes.

Conclusions: Baseline anxiety and depressive symptom severity may be relevant for understanding variability in clinically meaningful outcomes following brief psychological interventions in primary care. Patients presenting with greater depressive symptom burden may benefit from closer monitoring and more intensive or extended treatment pathways within stepped-care models.

Keywords

common mental disorders; primary care; symptom severity; reliable improvement; reliable recovery.

Introduction

Common mental disorders (CMDs) are among the most frequent reasons for consultation in primary care [1], affecting an estimated 20%–30% of patients. Psychologi-

Submitted: 10 June 2026 Revised: 17 July 2026 Accepted: 7 August 2026 Online First: 29 October 2026
*Corresponding author details: César González-Blanch, Mental Health Centre, University Hospital Marqués de Valdecilla, 39008 Santander, Spain; IDIVAL, Valdecilla Biomedical Research Institute, 39011 Santander, Spain; Department of Psychology, International University of La Rioja (UNIR), 26006 Logroño, Spain; Faculty of Health Sciences, European University of the Atlantic, 39011 Santander, Spain. Email: cesar.gonzalezblanch@scsalud.es

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

cal therapies delivered in primary care, particularly cognitive behavioural therapy (CBT), have consistently demonstrated greater effectiveness than usual care or no treatment in reducing symptoms of depression and anxiety, as supported by robust meta-analytic evidence [2–5]. However, access to evidence-based psychological treatments in primary care remains limited, with substantial treatment gaps, long waiting times, and an overreliance on pharmacological management across many healthcare systems [6–9].

In response to the growing need for accessible evidence-based treatments, several countries have expanded the provision of psychological therapies within primary care settings through initiatives such as National Health Service Talking Therapies in England and similar programmes in Norway and Australia [10–12]. In Spain, regional initiatives have also integrated clinical psychologists into primary care settings. Within this context, a regional primary care psychological intervention programme was implemented in Cantabria in 2023. First-year outcomes from the programme showed clinically meaningful reductions in depressive and anxiety symptoms, high rates of reliable improvement and recovery, and very low deterioration rates under routine care conditions [13].

Nevertheless, clinically meaningful outcomes vary considerably across patients receiving brief psychological interventions in routine primary care settings. Previous research has identified baseline symptom severity as an important correlate of psychological therapy outcomes, although findings have varied depending on the outcome examined [14,15]. Higher baseline symptom severity has been associated with poorer prognosis and lower likelihood of recovery in several naturalistic and primary care studies, particularly for depressive symptoms, which are linked to greater chronicity and residual impairment [14]. However, other studies suggest that patients with higher baseline severity may also show greater absolute symptom improvement following psychotherapy [16,17]. Notably, much of the existing literature has focused either on depression or anxiety separately, on disorder-specific samples, or on randomized controlled trials relying primarily on continuous symptom change scores [18]. This apparent inconsistency may partly reflect differences in how treatment outcomes are defined. Examining baseline depressive and anxiety severity in relation to both reliable improvement and reliable recovery is particularly relevant because these outcomes capture complementary dimensions of clinically meaningful change: symptom reduction beyond measurement error and the attainment of subclinical symptom levels. Such an approach may provide a more differentiated account of treatment outcome than analyses focused on a single symptom domain or on continuous change scores alone. Evidence from routine primary care psychological services in Spain remains limited, particularly for brief interventions delivered under real-world conditions.

The present study examined whether baseline depressive and anxiety symptom severity were associated with reliable improvement and reliable recovery in a real-world primary care psychological programme after adjustment for demographic variables and number of therapeutic sessions.

Materials and Methods

Study Design

The present study is a secondary analysis of data derived from a prospective, observational programme evaluating the implementation of a brief primary care–based psychological intervention for CMDs in the region of Cantabria, Spain, as described in detail elsewhere [13]. Data were collected from patients referred to The Clinical Psychology in Primary Care Programme between July 2023 and June 2024 across ten primary care centres: Puente San Miguel, Gama, Santoña, Colindres, Laredo, Saja-Cabuérniga in Cabezón de la Sal, Castilla-Hermida and Centro in Santander, and Tanos and Covadonga in Torrelavega.

Participants

In the present study, participants were included if they met all of the following criteria: (1) age 16 years or older; (2) CMD-related symptoms, specifically anxiety or depression, as the main reason for consultation; (3) scores at or above the established clinical thresholds, defined as $\geq$10 on the Generalized Anxiety Disorder-7 (GAD-7) and/or $\geq$10 on the Patient Health Questionnaire-9 (PHQ-9) [19,20]; (4) attendance at one or more individual treatment sessions; and (5) availability of both pre-treatment and post-treatment assessments. Exclusion criteria were the presence of a severe mental disorder; high suicide risk, including active suicidal ideation with planning; concurrent psychological treatment or specialised mental health care; medico-legal assessment; and intellectual disability.

Procedure

Entry into the programme was initiated through referral by the patient's general practitioner, after an initial clinical evaluation and psychometric screening. Eligibility was subsequently assessed by a clinical psychologist

Actas Esp Psiquiatr 2026;1–9. https://doi.org/10.62641/aep.2399 | ISSN:1578-2735
© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

during a comprehensive intake interview, using the predefined inclusion and exclusion criteria. Baseline demographic and clinical data were collected as part of routine care and recorded in the electronic health record. Age and sex were obtained from the clinical record. The PHQ-9 and GAD-7 were administered at two predefined time points: during the initial assessment, before treatment began, and at treatment completion. Both instruments assessed symptoms experienced during the previous two weeks and were administered as part of the programme's routine outcome-monitoring procedures. The number of treatment sessions was derived from routine attendance records. These data were subsequently extracted in pseudonymised form for the present analyses. After the initial assessment, patients received a diagnosis and an individual clinical formulation, which informed the tailoring of the intervention to their specific needs.

Intervention

The intervention was a brief psychological treatment provided either individually or in groups, with the format selected according to the treating clinical psychologist's clinical judgement. However, as group treatment, which represented fewer than 10% of cases, followed a fixed number of sessions for all participants, variability in treatment duration was limited to the individual format; therefore, only patients receiving individual treatment were included in the present analyses. The individual intervention began with a 60-minute assessment appointment and could include up to seven subsequent therapy sessions, each lasting 30–40 minutes, in which various evidence-based psychological techniques were applied.

Measures

Symptoms of Depression

Depressive symptoms were assessed using the PHQ-9 [19,21], a validated self-report instrument for screening and measuring depression severity. The scale includes nine items that reflect the diagnostic criteria for major depressive disorder in the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). Respondents indicate how often they have experienced each symptom during the previous two weeks using a 4-point Likert response format, ranging from 0 (not at all) to 3 (nearly every day). Scores of 10 or higher are commonly used to indicate clinically relevant depressive symptoms, a threshold that has shown 88% sensitivity and 88% specificity for major depressive disorder [19]. In the present sample, internal consistency at baseline was good (McDonald's $\omega = 0.86$).

Symptoms of Anxiety

Anxiety symptoms were measured with the GAD-7 [20,22], a widely used self-report measure of anxiety symptom severity and a screening tool for generalised anxiety disorder and other anxiety-related presentations. The GAD-7 contains seven items referring to symptoms experienced over the past two weeks, rated on a 4-point Likert scale from 0 (Not at all) to 3 (Nearly every day). A cut-off score of 10 is recommended to identify clinically significant anxiety, with reported sensitivity of 89% and specificity of 82% [20]. Baseline internal consistency in the present sample was good (McDonald's $\omega = 0.84$).

Data Analyses

Sociodemographic and clinical characteristics were summarised descriptively. For continuous variables, means and standard deviations (SDs) were calculated, whereas categorical variables were reported as counts and percentages. Following Jacobson and Truax's [23] method, the threshold for reliable change was calculated using the formula: $1.96 \times \mathrm{SD} \times \sqrt{2(1 - \mathrm{r})}$, where r represents the internal consistency and SD represents the standard deviation of baseline scores for the scale. Based on this calculation, reliable improvement was defined as a reduction of at least 7 points on the PHQ-9 or at least 6 points on the GAD-7. Participants were classified as reliably improved if they met the reliable-change criterion on at least one of the two measures. Reliable recovery was defined as reliable improvement on at least one measure together with post-treatment scores below the clinical cut-off (<10) on both the PHQ-9 and GAD-7. Participants were then classified according to whether they achieved reliable improvement and reliable recovery. Baseline demographic and clinical characteristics were compared between outcome groups using independent-samples *t*-tests for continuous variables and chi-square tests for categorical variables.

Reliable clinical improvement and reliable recovery were examined using binary logistic regression analyses. Associations between baseline demographic and clinical variables (age, sex, baseline PHQ-9 and GAD-7 scores, and number of sessions) and treatment outcomes were examined by entering all predictors simultaneously into each model. As a sensitivity analysis, both logistic regression models were repeated without number of sessions, because treatment duration was determined during therapy and may have been influenced by treatment response and

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

clinical decision-making. To assess whether the association between baseline symptom severity and reliable improvement was influenced by the mathematical properties of the reliable-change criterion, additional sensitivity analyses were conducted using continuous post-treatment PHQ-9 and GAD-7 scores as dependent variables. Each linear regression model included the corresponding baseline symptom score, the other baseline symptom measure, age, sex, and number of sessions as covariates. Assumptions for logistic regression were examined prior to analyses. Multicollinearity diagnostics indicated no relevant collinearity among predictors (all variance inflation factors [VIFs] <2). Linearity of the logit for continuous predictors was assessed using the Box–Tidwell procedure, and model fit was evaluated using the Hosmer–Lemeshow test. Results are presented as odds ratios (ORs) with 95% confidence intervals (CIs), whereas results from the linear regression models are reported as unstandardized coefficients (B), standard errors (SEs), and standardized coefficients (β). Statistical significance was set at $p < 0.05$ (two-tailed) for all analyses. All analyses were conducted using IBM SPSS Statistics for Windows, version 26 (IBM Corp., Armonk, NY, USA).

Results

Baseline Characteristics

Of the 1290 patients referred to the programme, 215 (16.7%) were excluded because they received a non-individual intervention or the intervention type was not recorded, 127 (9.8%) because baseline PHQ-9/GAD-7 data were missing or incomplete, 21 (1.6%) because session data were missing, 92 (7.1%) because no follow-up treatment sessions were recorded, and 323 (25.0%) because post-treatment PHQ-9/GAD-7 data were missing or incomplete. The final analytical sample therefore comprised 512 participants. Comparisons between included and non-included patients indicated no meaningful differences in sex or baseline depression and anxiety severity. Included participants were slightly older than non-included participants, although the effect size was small (d = 0.13).

Of these, 333 (65.0%) were women, with a mean age of 45.0 years (SD = 13.9). Baseline mean scores were 15.9 (SD = 5.3) for the PHQ-9 and 14.8 (SD = 3.8) for the GAD-7. Participants attended a mean of 4.3 sessions (SD = 1.3). Reliable improvement was observed in 73.6% of participants, whereas 53.3% achieved reliable recovery. Table 1 summarizes baseline demographic and clinical characteristics according to reliable improvement and reliable recovery status. Reliable improvement was associated with higher baseline PHQ-9 and GAD-7 scores, whereas reliable recovery was associated with lower baseline PHQ-9 scores and fewer treatment sessions. Age and sex were not significantly associated with outcomes.

Associations With Reliable Clinical Improvement and Reliable Recovery

Logistic regression models examining associations between baseline demographic and clinical variables and treatment outcomes (i.e., reliable improvement and reliable recovery) are presented in Table 2. Higher baseline anxiety symptom severity (GAD-7) was associated with a greater likelihood of reliable improvement, whereas a higher number of sessions was associated with a lower likelihood of reliable improvement; no other variables were significantly associated with this outcome. Regarding reliable recovery, lower baseline depressive symptom severity (PHQ-9) and fewer treatment sessions were associated with a greater likelihood of reliable recovery. Age, sex, and baseline anxiety symptom severity were not significantly associated with this outcome (Table 2). Sensitivity analyses excluding number of sessions yielded an essentially unchanged pattern of results (**Supplementary Table 1**). Base-

Table 1. Baseline demographic and clinical characteristics according to reliable improvement and reliable recovery status (n = 512).

Variable	No reliable improvement (n = 135)	Reliable improvement (n = 377)	Test statistic	*p*-values	No reliable recovery (n = 239)	Reliable recovery (n = 273)	Test statistic	*p*-values
Age, mean (SD)	46.1 (14.1)	44.6 (13.8)	$t = 1.08$	0.280	46.0 (13.43)	44.1 (14.3)	$t = 1.55$	0.121
Sex, female, n (%)	86 (63.7)	247 (65.5)	$\chi^2 = 0.14$	0.705	158 (66.1)	175 (64.1)	$\chi^2 = 0.23$	0.635
PHQ-9 baseline, mean (SD)	14.9 (5.5)	16.3 (5.2)	$t = -2.59$	0.010	17.0 (5.4)	15.0 (5.07)	$t = 4.23$	<0.001
GAD-7 baseline, mean (SD)	13.4 (4.1)	15.3 (3.6)	$t = -5.03$	<0.001	15.06 (4.0)	14.5 (3.67)	$t = 1.73$	0.085
Number of sessions, mean (SD)	4.4 (1.3)	4.2 (1.33)	$t = 1.10$	0.274	4.5 (1.3)	4.1 (1.4)	$t = 4.10$	<0.001

Note: SD, standard deviation; PHQ-9, Patient Health Questionnaire-9; GAD-7, Generalized Anxiety Disorder-7; *t*, independent-samples *t*-test statistic; χ^2, chi-square statistic.

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

Table 2. Logistic regression model examining associations of baseline demographic and clinical variables with reliable recovery and reliable improvement (n = 512).

Predictor	Reliable improvement OR (95% CI)	*p*-values	Reliable recovery OR (95% CI)	*p*-values
Sex, female	0.94 (0.62–1.43)	0.770	1.06 (0.72–1.54)	0.773
Age	0.99 (0.98–1.01)	0.390	0.99 (0.98–1.00)	0.066
Number of sessions	0.83 (0.71–0.98)	0.027	0.76 (0.66–0.88)	<0.001
PHQ-9 baseline	1.00 (0.95–1.04)	0.866	0.92 (0.88–0.96)	<0.001
GAD-7 baseline	1.16 (1.08–1.23)	<0.001	1.04 (0.98–1.10)	0.182

Note: OR, odds ratio; CI, confidence interval; PHQ-9, Patient Health Questionnaire-9; GAD-7, Generalized Anxiety Disorder-7.

line GAD-7 remained associated with reliable improvement (OR = 1.14, 95% CI: 1.07–1.22, $p < 0.001$), whereas baseline PHQ-9 remained inversely associated with reliable recovery (OR = 0.92, 95% CI: 0.88–0.96, $p < 0.001$). No other baseline variables were significantly associated with either outcome. However, sensitivity analyses using continuous post-treatment symptom scores did not reproduce the favourable association between higher baseline anxiety severity and reliable improvement (**Supplementary Table 2**). Higher baseline GAD-7 scores were associated with higher post-treatment GAD-7 scores (B = 0.20, SE = 0.06, $\beta = 0.16$, $p = 0.001$). Baseline PHQ-9 scores were associated with both higher post-treatment GAD-7 scores (B = 0.15, SE = 0.04, $\beta = .17$, $p = 0.001$) and higher post-treatment PHQ-9 scores (B = 0.44, SE = 0.05, $\beta = .42$, $p < 0.001$), whereas baseline GAD-7 was not associated with post-treatment PHQ-9 scores ($p = 0.420$).

Discussion

The present study examined associations between baseline depressive and anxiety symptom severity and clinically meaningful patient-level outcomes in a real-world primary care psychological intervention programme. Reliable improvement was observed in nearly three-quarters of participants, whereas more than half achieved reliable recovery following treatment. Higher baseline anxiety symptom severity was associated with a greater likelihood of reliable improvement, whereas lower baseline depressive symptom severity was associated with a greater likelihood of reliable recovery. In addition, a lower number of treatment sessions was associated with a greater likelihood of both reliable improvement and reliable recovery, while age and sex were not significantly associated with either outcome.

These findings suggest that baseline anxiety and depressive symptom severity have differential relationships with clinically meaningful treatment outcomes in brief psychological interventions delivered in primary care settings. One plausible explanation is that anxiety symptoms may be more responsive in the short term to brief psychological interventions. This interpretation is supported by meta-analytic evidence showing robust effects of CBT for anxiety disorders across settings [24,25], as well as by evidence indicating that early symptom change tends to be a stronger predictor of final outcomes in anxiety than in depression [26]. Similarly, the PsicAP study, a large primary care randomized controlled trial of brief transdiagnostic CBT conducted within the Spanish National Health System, yielded larger post-treatment effect sizes for anxiety than for depressive symptoms [27]. In addition, previous research suggests that patients with higher baseline symptom severity may experience greater absolute symptom change while still showing lower probabilities of achieving recovery or remission thresholds, particularly in depression [14,16]. It should nevertheless be acknowledged that reliable improvement partly depends on baseline symptom levels, because patients with greater initial severity have more opportunity to achieve the magnitude of change required to exceed reliable-change thresholds. Consistent with this concern, sensitivity analyses using continuous post-treatment symptom scores did not reproduce a favourable association between higher baseline anxiety severity and outcome. Instead, higher baseline GAD-7 scores were associated with greater residual anxiety symptoms at post-treatment. Baseline depressive severity showed a broader pattern, being associated with both post-treatment depressive and anxiety symptoms, whereas baseline anxiety severity was associated only with post-treatment anxiety. This asymmetric pattern may suggest that depressive symptom burden reflects a broader marker of persistent emotional distress, although the overlap between depressive and anxiety symptoms and the exploratory nature of these analyses warrant cautious interpretation. Accordingly, the association between baseline GAD-7 and reliable improvement should not be interpreted as evidence that greater initial anxiety severity confers a clinical advantage. On the other hand, moderate to severe depressive symptoms have been con-

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

sistently associated with greater chronicity, residual symptom persistence, and recurrent courses in longitudinal research [28–30]. Indeed, longitudinal trajectory studies indicate that depressive disorders are more likely to follow persistent or recurrent courses over time, whereas anxiety disorders show more heterogeneous and episodic patterns [31]. Similarly, large routine-care studies from Improving Access to Psychological Therapies (IAPT) services have shown that higher baseline symptom severity may predict faster symptomatic change while simultaneously reducing the likelihood of symptomatic recovery [32]. In line with this, Prieto-Vila *et al*. [33] found that, in primary care populations, depressive symptoms follow heterogeneous longitudinal trajectories, with poorer trajectories associated with higher baseline severity.

Although a lower number of treatment sessions was associated with a greater likelihood of reliable improvement and reliable recovery, this finding should be interpreted cautiously. Given the observational and naturalistic design of the study, treatment duration was determined by clinical judgment rather than experimental allocation. Therefore, the number of sessions is likely to reflect underlying clinical complexity, slower response trajectories, or persistent symptomatology, with patients experiencing greater difficulties or poorer early response being more likely to receive additional sessions. This interpretation is consistent with good-enough level models of psychotherapy, which propose that treatment duration is shaped by individual rates of improvement rather than by a simple cumulative dose effect [34–36]. Consequently, this association may partly reflect indication bias rather than a detrimental effect of longer treatment exposure itself, as treatment duration in observational studies is often associated with baseline severity or prognosis rather than treatment effectiveness [37]. Nevertheless, sensitivity analyses excluding number of sessions produced essentially the same associations between baseline symptom severity and both outcomes, indicating that the main findings were not dependent on the inclusion of treatment duration in the models.

These findings have important implications for primary care services facing high demand and limited resources. In particular, the results suggest that baseline depressive and anxiety symptom severity may be relevant to understanding different dimensions of treatment outcome, including reliable improvement and reliable recovery. The lower likelihood of reliable recovery among patients with more severe depressive symptoms raises the hypothesis that this subgroup may require closer monitoring or alternative treatment pathways, but this possibility should be evaluated in prospective studies directly comparing treatment intensity, duration, and sequencing. At the same time, the association between higher baseline anxiety severity and reliable improvement should be interpreted cautiously, as it was not reproduced when continuous post-treatment anxiety symptoms were examined. More broadly, these results highlight the importance of distinguishing between symptom improvement and recovery when evaluating psychotherapy outcomes, as different baseline clinical characteristics may be differentially associated with these patient-level outcomes.

This study has several strengths, including its large naturalistic primary care sample, the use of routinely collected data reflecting real-world clinical practice, and its focus on clinically meaningful outcomes such as reliable improvement and recovery, which are directly relevant to service evaluation. At the same time, several limitations should also be acknowledged. First, the study relied exclusively on pre- and post-treatment assessments, preventing the examination of symptom trajectories or early response patterns. Relatedly, the absence of follow-up assessments precludes conclusions regarding the long-term stability of reliable improvement and reliable recovery over time. Second, the analyses included a limited set of routinely available covariates. Relevant clinical, sociodemographic, and treatment-related factors were not systematically available and may have influenced the observed associations. Accordingly, the present analyses should not be interpreted as a comprehensive prognostic model. Future prospective studies should include a broader range of such variables and formally evaluate predictive performance using measures of discrimination, calibration, and internal validation. Third, the analyses included only participants who completed both baseline and post-treatment assessments, which may have introduced some risk of selection bias. However, comparisons between participants included and not included in the analytical sample indicated that the groups were broadly comparable in sex distribution and baseline depressive and anxiety symptom severity, with only a small difference in age. Moreover, sensitivity analyses from the programme evaluation suggested that the main findings remained robust under different assumptions regarding missing post-treatment data. Clinician-recorded dropout rates were also substantially lower than the proportion of participants lacking post-treatment assessments, suggesting that a considerable proportion of missing post-treatment data likely reflected failures in routine outcome collection rather than true treatment dropout [13]. Finally, the observational and naturalistic design of the study precludes causal inferences regarding the associations identified, particularly with respect to treatment duration, which was determined by clinical judgment rather than experimental allocation.

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

Conclusions

Brief psychological interventions delivered in routine primary care settings were associated with high rates of reliable improvement and recovery among patients with CMDs. Baseline depressive and anxiety symptom severity showed different associations with treatment outcomes. Greater depressive severity was associated with a lower likelihood of reliable recovery and with higher residual symptom levels at post-treatment. Although higher baseline anxiety severity was associated with reliable improvement, this finding should be interpreted cautiously because it was not reproduced in sensitivity analyses using continuous post-treatment symptoms. These findings reinforce the importance of distinguishing between reliable improvement and reliable recovery when evaluating outcomes in routine primary care psychological therapy.

Availability of Data and Materials

The data are available from the corresponding author on reasonable request.

Author Contributions

CGB: Conceptualization, Methodology, Formal analysis, Writing—Original Draft, Writing—Review & Editing. MCS, RBS, CMN, DGS, SPM, VLS: Investigation, Data curation, Writing—Review & Editing. PDG: Project administration, Writing—Review & Editing. AP: Conceptualization, Methodology, Project administration, Data curation, Writing—Review & Editing. All authors read and approved the final manuscript. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Ethics Approval and Consent to Participate

The research protocol was reviewed and approved by the Clinical Research Ethics Committee of Cantabria (CEIm Cantabria; reference number: 2024.204). The study was conducted in accordance with the principles of the Declaration of Helsinki. The requirement for individual informed consent was waived by the ethics committee because the study used routinely collected data from a clinical care programme to evaluate its effectiveness.

Acknowledgment

We thank all the general practitioners, nurses, and administrative staff of the participating primary care centres for their collaboration in referral processes and programme implementation. We are particularly grateful to Dr. Óscar Fernández Torre for his instrumental role in the initiation of the programme. We also acknowledge the institutional support of the Servicio Cántabro de Salud, which enabled access to the clinical data used in this study.

Funding

This work was supported by the Valdecilla Biomedical Research Institute (IDIVAL) under Grant PRIMVAL25/04.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.62641/aep.2399.

References

[1] Caspi A, Houts RM, Moffitt TE, Richmond-Rakerd LS, Hanna MR, Sunde HF, *et al*. A nationwide analysis of 350 million patient encounters reveals a high volume of mental-health conditions in primary care. Nature Mental Health. 2024; 2: 1208–1216. https://doi.org/10.1038/s44220-024-00310-5.

[2] Cape J, Whittington C, Buszewicz M, Wallace P, Underwood L. Brief psychological therapies for anxiety and depression in primary care: meta-analysis and meta-regression. BMC medicine. 2010; 8: 38. https://doi.org/10.1186/1741-7015-8-38.

[3] Zhang A, Franklin C, Jing S, Bornheimer LA, Hai AH, Himle JA, *et al*. The effectiveness of four empirically supported psychotherapies for primary care depression and anxiety: A systematic review and meta-analysis. Journal of affective disorders. 2019; 245: 1168–1186. https://doi.org/10.1016/j.jad.2018.12.008.

[4] Zhang A, Borhneimer LA, Weaver A, Franklin C, Hai AH, Guz S, *et al*. Cognitive behavioral therapy for primary care depression and anxiety: a secondary meta-analytic review using robust variance estimation in meta-regression. Journal of behavioral medicine. 2019; 42: 1117–1141. https://doi.org/10.1007/s10865-019-00046-z.

[5] Twomey C, O'Reilly G, Byrne M. Effectiveness of cognitive behavioural therapy for anxiety and depression in primary care: a meta-analysis. Family practice. 2015; 32: 3–15.

[6] Thornicroft G, Chatterji S, Evans-Lacko S, Gruber M, Sampson N,

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

Aguilar-Gaxiola S, *et al*. Undertreatment of people with major depressive disorder in 21 countries. British Journal of Psychiatry. 2017; 210: 119–124. https://doi.org/10.1192/bjp.bp.116.188078.

[7] Freeman M. The World Mental Health Report: transforming mental health for all. World Psychiatry. 2022; 21: 391–392. https://doi.org/10.1002/wps.21018.

[8] Kazdin AE. Addressing the treatment gap: Expanding the scalability and reach of treatment. Journal of consulting and clinical psychology. 2023; 91: 3–5. https://doi.org/10.1037/ccp0000762.

[9] Miras-Aguilar MDM, Ruiz-Gutierrez J, Martinez-Gomez S, Perez-Garcia-Abad S, Ramos-Barron C, Pariente-Rodrigo E, *et al*. Waiting Times in Clinical Psychology in Public Mental Health Units: Predictors of Attendance at the First Appointment and Early Dropout. Psicothema. 2026; 38: 13–22. https://doi.org/10.70478/psicothema.2026.38.02.

[10] Cromarty P, Drummond A, Francis T, Watson J, Battersby M. NewAccess for depression and anxiety: adapting the UK Improving Access to Psychological Therapies Program across Australia. Australasian psychiatry. 2016; 24: 489–492. https://doi.org/10.1177/1039856216641310.

[11] Knapstad M, Lervik LV, Saether SMM, Aaro LE, Smith ORF. Effectiveness of Prompt Mental Health Care, the Norwegian Version of Improving Access to Psychological Therapies: A Randomized Controlled Trial. Psychotherapy and psychosomatics. 2020; 89: 90–105. https://doi.org/10.1159/000504453.

[12] Wakefield S, Kellett S, Simmonds-Buckley M, Stockton D, Bradbury A, Delgadillo J. Improving Access to Psychological Therapies (IAPT) in the United Kingdom: A systematic review and meta-analysis of 10-years of practice-based evidence. The British journal of clinical psychology. 2021; 60: 1–37. https://doi.org/10.1111/bjc.12259.

[13] González-Blanch C, Rodríguez-Perez N, Camino-Sánchez M, Bengochea-Seco R, Montes-Novoa C, Gil-Sanz D, *et al*. Real-world effectiveness of integrating clinical psychologists into primary care: First-year outcomes from a regional programme for common mental disorders in Spain. The British journal of clinical psychology. 2026. https://doi.org/10.1111/bjc.70068. (Online ahead of print)

[14] Buckman JEJ, Saunders R, Cohen ZD, Barnett P, Clarke K, Ambler G, *et al*. The contribution of depressive 'disorder characteristics' to determinations of prognosis for adults with depression: an individual patient data meta-analysis. Psychological Medicine. 2021; 51: 1068–81. https://doi.org/10.1017/S0033291721001367.

[15] González-Blanch C, Munoz-Navarro R, Medrano LA, Moriana JA, Ruiz-Rodriguez P, Cano-Vindel A. Moderators and predictors of treatment outcome in transdiagnostic group cognitive-behavioral therapy for primary care patients with emotional disorders. Depression and anxiety. 2021; 38: 757–767. https://doi.org/10.1002/da.23164.

[16] Cuijpers P, Harrer M, Miguel C, Karyotaki E, Papola D. Does baseline severity interact with the effects of psychotherapy for depression? A meta-analytic review. Journal of affective disorders. 2026; 399: 121106. https://doi.org/10.1016/j.jad.2025.121106.

[17] Mohr DC, Kwasny MJ, Meyerhoff J, Graham AK, Lattie EG. The effect of depression and anxiety symptom severity on clinical outcomes and app use in digital mental health treatments: Meta-regression of three trials. Behaviour research and therapy. 2021; 147: 103972. https://doi.org/10.1016/j.brat.2021.103972.

[18] Spinhoven P, Kraaij V, Garnefski N, Witlox M. Predictors of Outcome of Acceptance and Commitment Therapy and Cognitive Behavioral Therapy in Older Adults With Anxiety Symptoms. Psychological report. 2025: 332941251399176. https://doi.org/10.1177/00332941251399176. (Online ahead of print)

[19] Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. Journal of general internal medicine. 2001; 16: 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

[20] Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Archives of internal medicine. 2006; 166: 1092–1097. https://doi.org/10.1001/archinte.166.10.1092.

[21] Diez-Quevedo C, Rangil T, Sanchez-Planell L, Kroenke K, Spitzer RL. Validation and utility of the patient health questionnaire in diagnosing mental disorders in 1003 general hospital Spanish inpatients. Psychosomatic medicine. 2001; 63: 679–686. https://doi.org/10.1097/00006842-200107000-00021.

[22] Garcia-Campayo J, Zamorano E, Ruiz MA, Pardo A, Perez-Paramo M, Lopez-Gomez V, *et al*. Cultural adaptation into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. Health and quality of life outcomes. 2010; 8: 8. https://doi.org/10.1186/1477-7525-8-8.

[23] Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. Journal of consulting and clinical psychology. 1991; 59: 12–19. https://doi.org/10.1037//0022-006x.59.1.12.

[24] Hofmann SG, Asnaani A, Vonk IJ, Sawyer AT, Fang A. The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses. Cognitive therapy and research. 2012; 36: 427–440. https://doi.org/10.1007/s10608-012-9476-1.

[25] Cuijpers P, Cristea IA, Karyotaki E, Reijnders M, Huibers MJ. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. World psychiatry. 2016; 15: 245–258. https://doi.org/10.1002/wps.20346.

[26] Beard JIL, Delgadillo J. Early response to psychological therapy as a predictor of depression and anxiety treatment outcomes: A systematic review and meta-analysis. Depression and anxiety. 2019; 36: 866–878. https://doi.org/10.1002/da.22931.

[27] Cano-Vindel A, Muñoz-Navarro R, Moriana JA, Ruiz-Rodríguez P, Medrano LA, González-Blanch C. Transdiagnostic group cognitive behavioural therapy for emotional disorders in primary care: the results of the PsicAP randomized controlled trial. Psychological medicine. 2022; 52: 3336–3348. https://doi.org/10.1017/S0033291720005498.

[28] Bockting CL, Hollon SD, Jarrett RB, Kuyken W, Dobson K. A lifetime approach to major depressive disorder: The contributions of psychological interventions in preventing relapse and recurrence. Clinical psychology review. 2015; 41: 16–26. https://doi.org/10.1016/j.cpr.2015.02.003.

[29] Judd LL, Paulus MJ, Schettler PJ, Akiskal HS, Endicott J, Leon AC, *et al*. Does incomplete recovery from first lifetime major depressive episode herald a chronic course of illness? The American journal of psychiatry. 2000; 157: 1501–1504. https://doi.org/10.1176/appi.ajp.157.9.1501.

[30] Hardeveld F, Spijker J, De Graaf R, Nolen WA, Beekman AT. Recurrence of major depressive disorder and its predictors in the gen-

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

eral population: results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). Psychological medicine. 2013; 43: 39–48. https://doi.org/10.1017/S0033291712002395.

[31] Olino TM, Klein DN, Lewinsohn PM, Rohde P, Seeley JR. Latent trajectory classes of depressive and anxiety disorders from adolescence to adulthood: descriptions of classes and associations with risk factors. Comprehensive psychiatry. 2010; 51: 224–235. https://doi.org/10.1016/j.comppsych.2009.07.002.

[32] Stochl J, Soneson E, Stuart F, Fritz J, Walsh AEL, Croudace T, *et al*. Determinants of patient-reported outcome trajectories and symptomatic recovery in Improving Access to Psychological Therapies (IAPT) services. Psychological medicine. 2022; 52: 3231–3240. https://doi.org/10.1017/S0033291720005395.

[33] Prieto-Vila M, González-Blanch C, Estupiñá Puig FJ, Buckman JEJ, Saunders R, Muñoz-Navarro R, *et al*. Long-term depressive symptom trajectories and related baseline characteristics in primary care patients: Analysis of the PsicAP clinical trial. European psychiatry. 2024; 67: e32. https://doi.org/10.1192/j.eurpsy.2024.27.

[34] Baldwin SA, Berkeljon A, Atkins DC, Olsen JA, Nielsen SL. Rates of change in naturalistic psychotherapy: contrasting dose-effect and good-enough level models of change. Journal of consulting and clinical psychology. 2009; 77: 203–211. https://doi.org/10.1037/a0015235.

[35] Stiles WB, Barkham M, Connell J, Mellor-Clark J. Responsive regulation of treatment duration in routine practice in United Kingdom primary care settings: replication in a larger sample. Journal of consulting and clinical psychology. 2008; 76: 298–305. https://doi.org/10.1037/0022-006X.76.2.298.

[36] Stulz N, Lutz W, Kopta SM, Minami T, Saunders SM. Dose-effect relationship in routine outpatient psychotherapy: does treatment duration matter? Journal of counseling psychology. 2013; 60: 593–600. https://doi.org/10.1037/a0033589.

[37] Freemantle N, Marston L, Walters K, Wood J, Reynolds MR, Petersen I. Making inferences on treatment effects from real world data: propensity scores, confounding by indication, and other perils for the unwary in observational research. BMJ. 2013; 347: f6409. https://doi.org/10.1136/bmj.f6409.

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.