



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/245406/>

Version: Published Version

Article:

Malins, S., Saxon, D., Figueredo, G. et al. (2026) Predicting patient outcomes from initial in-session interactions within cognitive behavioral therapy and person-centered experiential therapy for severe depression. Behaviour Research and Therapy, 205. 105152. ISSN: 0005-7967

<https://doi.org/10.1016/j.brat.2026.105152>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. 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.



Predicting patient outcomes from initial in-session interactions within cognitive behavioral therapy and person-centered experiential therapy for severe depression

Sam Malins^{a,b,c,*}, David Saxon^d, Graziela Figueredo^e, Jeremie Clos^e, Thomas Trimble^e, Kavan Fatehi^f, David Waldram^b, Naomi Boycott^b, Alexandra Frost^a, Christian Sales^a, Kate Horton^{a,b}, Gillian E. Hardy^d, Michael Barkham^d, Richard Morriss^{a,b,c}, Thomas Schröder^{a,b,c}, Nima Moghaddam^g

^a Nottinghamshire Healthcare NHS Foundation Trust, United Kingdom

^b Institute of Mental Health, Nottingham, United Kingdom

^c University of Nottingham, School of Medicine, United Kingdom

^d University of Sheffield, School of Psychology, United Kingdom

^e University of Nottingham, School of Computer Science, United Kingdom

^f King's College London, Institute of Psychiatry, Psychology & Neuroscience, United Kingdom

^g University of Lincoln, School of Psychology Sport Science & Wellbeing, United Kingdom

ARTICLE INFO

Keywords:

Patient activation
Cognitive behavioral therapy
Person-centered experiential therapy
Consultation interactions coding scheme
Depression

ABSTRACT

Objective: Cognitive behavioral therapy (CBT) and person-centered experiential therapy (PCET) are the most frequently delivered individual psychological therapies in the UK National Health Service Talking Therapies program. However, many patients do not benefit from treatment and motivational processes early in therapy may influence outcomes. This study assessed specific, in-session interaction-types at the first session of CBT and PCET for severe depression that may predict therapeutic outcomes.

Method: Trained raters assessed active engagement in specific interactions from 103 first session transcripts of CBT ($n = 40$) and PCET ($n = 63$) for severe depression from a non-inferiority, pragmatic randomized controlled trial. Multilevel modelling and logistic regressions were used to evaluate whether initial patient-therapist interactions predicted depression change across therapy.

Results: In both CBT and PCET, patients' engagement with therapy-related action planning at initial sessions predicted greater improvement in depression. Conversely, where more time was taken with administrative aspects of completing measures (as opposed to discussing the implications of results) at first sessions, poorer outcomes ensued in both therapies. In CBT alone, engagement in identifying protective factors at session one predicted greater outcome improvement. In PCET alone, where more of session one was spent discussing patients' problems and contexts, greater outcome improvements followed.

Conclusion: Specific interaction-types at first sessions of CBT and PCET for severe depression may predict outcomes, with key similarities and differences. Initial sessions cannot be considered solely information gathering and assessment processes. Identification and deliberate practice of prognostic interactions could improve therapeutic skills and outcomes from early sessions.

1. Introduction

Depressive disorders are common and disabling conditions that are one of the leading causes of disability worldwide (Rong et al., 2025). Cognitive behavioral therapy (CBT; Beck, 2020) and person-centered

experiential therapy (PCET; Murphy, 2019) are the two most frequently delivered individual psychological therapies delivered in the UK National Health Service (NHS) Talking Therapies program psychological therapies; they are similarly effective in routine care, despite having different active therapeutic processes (NHS Digital, 2025).

* Corresponding author. Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU, United Kingdom.

E-mail address: sam.malins@nottingham.ac.uk (S. Malins).

<https://doi.org/10.1016/j.brat.2026.105152>

Received 16 June 2026; Received in revised form 3 September 2026; Accepted 5 September 2026

Available online 6 September 2026

0005-7967/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Although both CBT and PCET are effective, many people seeking treatment do not experience improvement, with approximately 40% of patients reporting non-improvement in the English NHS Talking Therapies program (NHS Digital, 2025). Therapeutic effectiveness can be improved using the established method of Routine Outcome Monitoring (ROM): routinely collecting outcome measures, gaining and sharing feedback from this, then adapting the process or focus of therapy as a result (Barkham et al., 2023). However, ROM's impact is limited during initial sessions, as the algorithm-based outcome prediction models employed are less accurate at that point and 20-30 completed cases are recommended to reliably identify individualized skill deficits among therapists (Lutz et al., 2021). This is important because therapeutic improvements can begin very early in therapy, and those who report early change experience the greatest improvements from therapy overall (Aafjes-van Doorn & Sweeney, 2019). Such improvements may be related to enhanced therapeutic alliance, engagement and motivational processes that can occur even in the first sessions of therapy, despite this session typically being characterized as serving assessment and information gathering functions (Gülüm et al., 2025).

It is well-established that baseline variables predict therapeutic outcomes in psychological therapies for depression; baseline symptom severity, in particular, is shown to be a robust predictor of prognosis (Hodsoll et al., 2024; Kamardina et al., 2026). There is also growing interest in the predictive power of early in-session conversation dynamics, with specific therapist and patient utterances associated with post-therapy outcomes (Ewbank, Cummins, Tablan, Bateup, et al., 2020; Ewbank, Cummins, Tablan, Catarino, et al., 2020). However, few in-session coding assessments account for the interactive nature of therapist-client communications.

To address this issue, specific, interaction-types at initial therapy sessions have been identified as predictive of therapeutic outcomes (Malins et al., 2021a). The Consultation Interactions Coding Scheme (CICS; Malins, Moghaddam, et al., 2026) has been used in research to detect processes related to alliance, engagement, and motivation that can contribute to early improvements (see Method for fuller description). Overall, the CICS assesses patient activation at each interaction between patient and therapist. Patient activation is an individual's confidence and perceived ability to manage their own health (Hibbard et al., 2004; Malins et al., 2020). The CICS also detected significant changes in the interactions between patient and therapist in the session directly prior to sudden gains – large, sustained therapeutic improvements occurring between two sessions that predict greater long-term improvements (Tang & DeRubeis, 1999).

Previous research using CICS-rated early session interactions showed superior predictive validity compared to patient- or therapist-rated motivation and therapeutic alliance rated at the same session. Specifically, where patients were actively engaged in interactions about noticing beneficial changes from initial therapy sessions, greater improvements ensued across a range of outcomes. The same was true of interactions where an understanding of their problems was developed. Conversely, where larger proportions of early sessions were spent describing problems in CBT for health anxiety, poorer outcomes followed (Malins et al., 2021b).

Such in-session predictors of outcome are of great value in helping therapists improve their overall effectiveness and become more responsive to patients' needs earlier in therapy, particularly when therapists do not otherwise become more effective with time or experience alone, on average (Goldberg, Babins-Wagner et al., 2016; Goldberg, Rousmaniere, et al., 2016). Personalized therapist feedback, which includes in-session interactions that either help or hinder therapeutic progress, could substantially advance the emerging approach of deliberate practice. Deliberate practice involves receiving feedback on individualized therapeutic skill deficits from observations of therapeutic practice, then iteratively rehearsing and improving the identified skills (Miller et al., 2023). Deliberate practice can help therapists improve their therapeutic skills and clinical effectiveness (Diamond et al., 2025).

By identifying specific skills requiring practice in particular clinical contexts, such feedback may better target therapists' learning efforts (Malins, Figueredo, et al., 2026; Miller & Hubble, 2023).

However, previous studies of the CICS have only involved participants receiving CBT in specialist clinical settings, where the primary problems were anxiety disorders. This limits the generalizability of the CICS with other therapeutic modalities, generalist settings, and where depressive disorders are the primary problem. There may also be important differences in the types of interactions that predict outcomes in non-CBT therapies, such as PCET, or in the treatment of depression as opposed to anxiety. Therefore, the current study aimed to evaluate the predictive validity of the CICS in initial sessions of both CBT and PCET as delivered in primary care in an English clinical organization for patients presenting with severe depression. Severe depression was the focus of this study, because the greatest differential effects between CBT and PCET has been among participants with more severe depression (Barkham et al., 2021). Severe depression also presents important, challenging, clinical differences from more moderate depression. These factors include greater risk of relapse, greater chance of non-response to treatment, and poorer physical health among those with more severe depression (Frank et al., 2023; McIntyre et al., 2023; Moriarty et al., 2024).

The preregistered primary hypotheses for this study were, first, that CICS ratings of first sessions will predict rate of change in depression across sessions (using the Patient Health Questionnaire 9-items: PHQ-9) and, that CICS ratings of first sessions will predict reliable recovery from depression at the last session attended. Specifically, that greater patient activation observed within session interactions would be associated with greater therapeutic improvement, in similar patterns to those observed in previous studies involving the CICS.

2. Method

2.1. Design

An observational, predictive modelling design was applied to individual participant data from a pragmatic, non-inferiority, randomized controlled trial of CBT versus PCET for primary care patients meeting a diagnosis of moderate or severe depression (PRaCTICED; Barkham et al., 2021); only participants meeting diagnosis of severe depression were included in this study, as detailed below. The trial protocol was published prior to trial completion (Saxon et al., 2017) and the current study was pre-registered prior to data linking and analysis (MASKED REFERENCE; see [Supplementary Table 1](#) for the Strengthening the Reporting of Observational Studies in Epidemiology [STROBE] checklist). We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

2.2. Setting and participants

The trial was embedded in a routine NHS Talking Therapies organization – an English primary care psychological therapy service for common mental health problems, implemented nationally. The source trial was conducted in Sheffield NHS Talking Therapies organization (formerly Improving Access to Psychological Therapies) with 510 participants aged 18 years or older who sought treatment from the clinical organization. Potential participants were given trial information and invited for further assessment if they reported moderate depressive symptoms on the Patient Health Questionnaire 9-items (PHQ-9; Kroenke et al., 2001), moderate depression defined as a score of 12 or above), confirmed that depression was their primary concern, and had no strong preference for either CBT ($n = 256$) or PCET ($n = 254$). Patients with an organic condition, previous diagnosis of personality disorder, bipolar disorder, schizophrenia, or a long-term physical health condition were excluded. Eligible participants were randomly assigned (1:1) to receive up to 20 sessions of either CBT or PCET. Assessors were blind to

allocations.

Of 510 trial participants, 103 (40 CBT, 63 PCET) were included in this study having been assessed as severely depressed at baseline with the Clinical Interview Schedule-Revised (CIS-R; Lewis, 1994); remained within the clinical range for symptoms of depression at session one (PHQ-9 score ≥ 10 ; Kroenke et al., 2001); did not switch treatments between CBT and PCET; attended at least two sessions of therapy; consented to their therapy sessions being audio-recorded; had at least one of their first two sessions recorded with sufficient clarity of recording for transcription, and completed session-by-session outcome assessments. Included participants were seen by 28 therapists (16 CBT, 12 PCET) who treated between 1 and 18 participants each.

The selection approach was employed to focus the analysis on patients with more severe symptoms because therapist effects are greater in this patient group (Johns et al., 2018). The approach also removed participants who significantly improved prior to starting therapy, for greater confidence that therapeutic input may have contributed to any subsequent symptom change. The small number of participants who received elements of both therapies (i.e., switched treatments) were excluded, to focus analysis on the unique features of each therapy. In addition, the selection process ensured there was sufficient session content and outcome data to enable robust analyses; specifically, potential participants were excluded where target session recordings were unavailable due to participants declining, the audio recording being unclear, where recordings ended part way through, or where therapists did not record the session for other unknown reasons (see Fig. 1 for participant flowchart).

2.3. Sample size

This is a secondary analysis of existing data; therefore, no a priori

sample size calculation was made. However, change in level 2 variables (participants) was the central focus of this analysis and, at this level, a sample size of 103 meets the minimum sample size for the analyses conducted, based on recommendations from MLM simulation studies (Snijders & Bosker, 2012).

2.4. Measures

2.4.1. Demographics

Baseline demographics were collected by independent researchers blind to treatment-arm allocation. These included: age, gender, employment status, ethnicity and Index of Multiple Deprivation. Each of these variables were separately tested in the analytic models and significant predictors were retained in final models.

2.4.2. Patient Health Questionnaire 9-items (PHQ-9)

The PHQ-9 is an established self-report measure for symptoms of depression based on the diagnostic criteria for a major depressive episode; it has excellent test-retest reliability and good convergent validity with assessments of general health (Kroenke et al., 2001). The PHQ-9 was completed by participants at each therapy session and is the outcome of interest for this study. Reliable improvement and reliable recovery were used as indices of participant benefit from therapy. Reliable improvement on the PHQ-9 is deemed to be achieved where participants report a reduction in PHQ-9 score of 6-points or more from baseline to last treatment session (NHS Digital, 2025). Where a 6-point improvement also includes movement from the clinical range (≥ 10) to the sub-clinical range on the PHQ-9 (< 10), it is classified as achieving reliable recovery, aligned to national guidance (Clark & Oates, 2014).

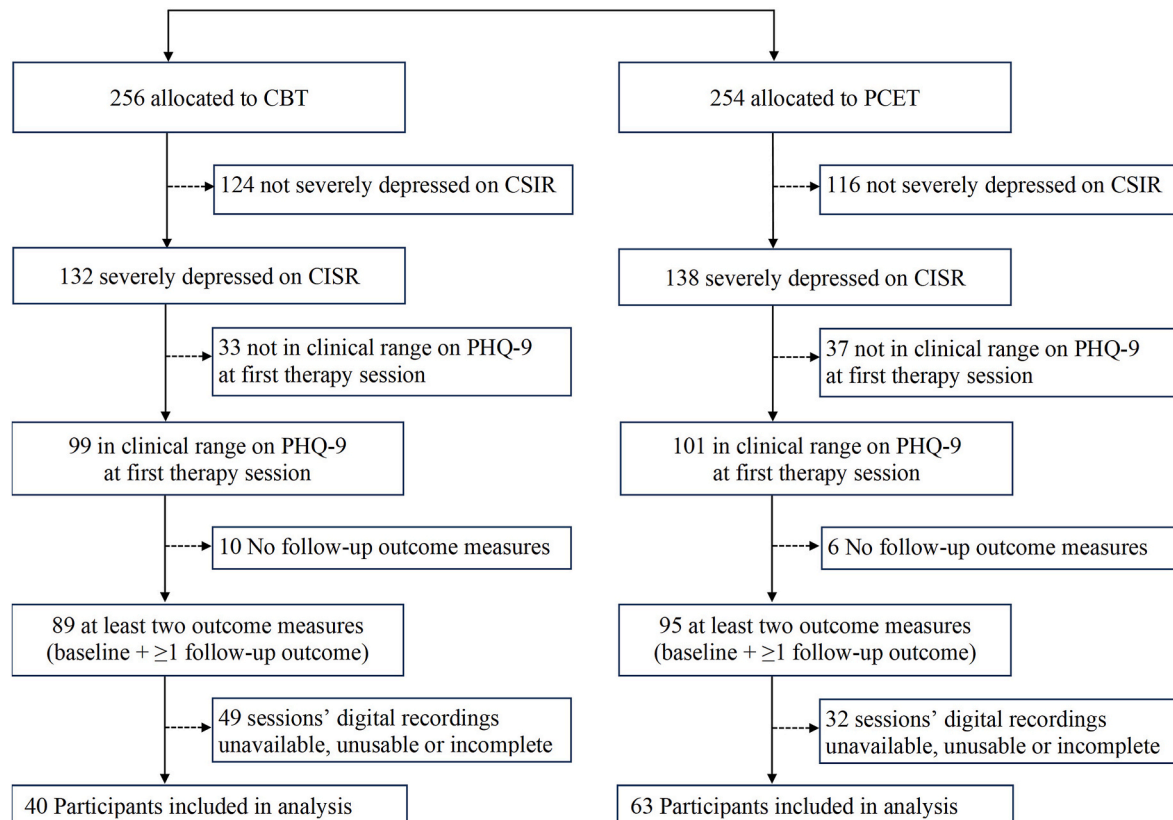


Fig. 1. Participant Flowchart.

Note. Abbreviations: Cbt, cognitive behavioral therapy; PCET, person-centered experiential therapy; CISR, clinical interview schedule revised; PHQ-9, patient health questionnaire 9-item.

2.4.3. Consultation interactions coding scheme (CICS)

The CICS classifies each in-session turn-of-speech on the interaction theme and level of patient activation identified, using transcripts rated by trained coders. Interaction themes are assigned using observable criteria for the following categories (Malins, Moghaddam, et al., 2026):

- **Action Planning and Idea generation (API):** Discussion of specific or potential plans for therapeutic activities outside the session.
- **Evaluations of Self or Therapy (EST):** Personal evaluations of the therapy process or of either participant in the therapeutic relationship.
- **Identifying Protective Factors (IPF):** Where pre-existing protective or helpful factors, unrelated to therapeutic input, are either identified or enquired about.
- **Information Discussion (IDI):** The giving, receiving, or requesting of specific information.
- **Measurement Completion (MC):** Instances when items from assessments are read aloud and/or responses given in order to complete such measures.
- **Noticing Change or Otherwise (NCO):** Reports of therapeutic change associated with therapeutic activities, or of a lack of change despite attempts to achieve it.
- **Other (OTH):** Interactions unrelated to therapy, such as discussion of parking practicalities or arranging dates and times of subsequent sessions.
- **Problem Analysis and Understanding (PAU):** The provision, receipt, or co-construction of an explanation or understanding of therapy-related problems.
- **Problem or Context Description (PCD):** Descriptions of problems or the contextual factors surrounding them.
- **Structuring and Task Focus (STF):** Verbal efforts to structure, plan, or progress the session, as well as instances in which the session deviates from therapy-related topics without intervention from either party.

After assigning an interaction theme, two coders rated the level of patient activation within each interaction. Activation scores ranged from +2 (observable, high levels of patient engagement and activation) to −2 (observable, low levels of activation and disengagement), with a score of 0 for interactions rated as neutral or unclear. These ratings corresponded to four established levels of patient activation plus the neutral category (see Table 1 for example descriptors for the Identifying Protective Factors interaction theme).

Session scores were computed by multiplying the percentage of transcribed words coded at each activation level within each interaction theme by the numerical value of that level. For example, if 10% of a transcript was coded as API at +2, the contribution to the API total would be $10\% \times (+2) = 20$; if a further 10% was coded as API at −1, the contribution would be $10\% \times (-1) = -10$, yielding a net API score of 10 for that session. This procedure was applied across all themes and transcripts; themes that were not present in a transcript were assigned a score of zero. Three interaction themes were rated as present-absent: Problem and Context Description, Measurement Completion, and Other interactions. For these interaction themes, the overall session score was based on the percentage of the transcript where these themes were rated as present.

2.5. Interventions

2.5.1. Cognitive behavioral therapy

Cognitive behavioral therapy (CBT) is an established, evidence-based treatment for a range of mental health problems and is recommended as a first line treatment for depression (National Institute for Health and Care Excellence, 2022). Beckian Cognitive Therapy (Beck, 2020) was delivered by trained high-intensity CBT practitioners in accordance with NICE guidance, which recommends CBT for severe

Table 1
Identifying Protective Factors CICS theme level descriptors and exemplar quotes.

Level	Level Descriptor ^a	Exemplar Quote
+2	<i>Patient initiated and identified protective factors that pre-date the onset of therapy, but the benefits of which are still being felt.</i>	P: I think that's good to be able to talk to people about. I do sometimes talk to, you know, my sister or my brother sometimes, we can talk about stuff or talk to a friend about it, so I think that helps to get it out rather than keep it in. T: Yes, or bottle it up, yes, I think you're right. [P 250]
+1	Therapist initiated or prompted pre-existing protective factors are described. The patient <i>strongly agrees</i> or identifies factors with <i>development, summary or corrections</i> .	T: Is there anything else you're doing at the moment that you think's helped? P: Like the band's always helped to be fair and writing music and – that's just the main thing that I feel like I'm living for at the minute. Just enjoying time with my mates and just like playing gigs and seeing what road it takes us down. [P 075]
0	Therapist initiated or prompted pre-existing protective factors – <i>no development or clear sense of patient view</i> OR very minor acknowledgment of change by the patient not personally “owned”.	T: Right and so going back to realizing I do have friends and they do care and I can talk to them and I've got roots []. P: Yeah. [P 518]
−1	Therapist or patient-identified protective factor with <i>unresolved negativity, ambivalence or minor rejection</i> of the protective factor by the patient or therapist.	T: I was wondering whether you felt a sense of, erm, achievement by doing that cos it sounds like it was really hard for you to do it and yet you did it. P: Erm, no. I don't think so. [P 142]
−2	Therapist or patient initiated protective factor – <i>ignored, refuted or distracted to another theme</i> and no agreement is reached on any beneficial impact.	T: So, it seems like your Dad has been a really helpful part of your life and his input is stopping things getting worse. P: What? That couldn't be farther from the truth. He is the main cause of my problems if anything [hypothesized]

Note. ^aItalics add emphasis to the key component of the level descriptor; Abbreviations: P, Patient; T, Therapist.

depression. CBT delivery was standardized using a core text and a CBT Clinical Practice Guide that was adapted from the COBRA trial (Beck, 2020; Rhodes et al., 2014).

2.5.2. Person-centered experiential therapy

Person-Centered Experiential Therapy (PCET) is developed from the core competences required to deliver effective humanistic treatments for depression and applies humanistic approaches with the strongest evidence from controlled trials (Elliott et al., 2021). PCET integrates principles from person-centered counseling and emotion-focused therapy, which share theoretical and methodological foundations (Mearns & Thorne, 2007). The PCET competencies are summarized in a framework informed by studies and key texts identified by the Humanistic Psychological Therapies Expert Reference Group (Hill, 2011). PCET delivery was standardized using a core text used with a treatment manual and associated resources (Murphy, 2019).

2.5.3. Treatment delivery

All treating therapists in both arms had to meet specified levels of training prior to being active on the trial and received appropriate training and regular supervision according to nationally mandated standards. Experienced, national trainers assessed treatment fidelity in a subsample of cases in the original trial using adherence scales designed specifically for each treatment modality (i.e., Cognitive Therapy Scale-Revised, Person-Centered Experiential Psychotherapy Scale).

Participants in both conditions were typically offered 20 sessions, though some cases exceeded this limit.

2.6. Procedure

Digital recordings of the first¹ therapy session received by participants through the trial were transcribed verbatim and anonymized. Each session was then coded with the CICS by two trained raters, using NVivo 12. Raters were three clinical psychologists, a researcher and an assistant psychologist all trained using a published manual (Malins, Moghaddam, et al., 2026). Training involved trainees reading the coding manual and attending three, weekly, 2-h training sessions, where transcripts were jointly live coded between an experienced coder and the trainee. The trainee coded two benchmarked transcripts independently between each training session, then discrepancies in codes from the benchmark ratings were addressed and agreed by discussion. Ratings were conducted independently, and raters could not access other raters' transcripts until all assessments were completed. Raters were blind to outcome and arm allocation. Inter-rater reliability was assessed at the level of turn-by-turn ratings and overall CICS scores.

2.7. Method of analysis

Inter-rater reliability was assessed for CICS ratings using two-way, random effects, consistency intra-class correlation coefficients (ICCs). In IRR, two CICS themes achieved ICCs below 0.5: Problem Analysis and Understanding (ICC = 0.20) and Noticing Change or Otherwise (ICC = 0.19). Ratings for these themes were removed from predictive modelling. The remaining CICS themes achieved an mean ICC of 0.73, indicating good IRR (Cicchetti, 1994; Supplementary Table 2). Coding proportions were calculated by dividing the number of words rated at each CICS theme in each transcript, by the total transcript words.

Multilevel Models (MLM) were used to assess predictors of PHQ-9 change over time, due to significant clustering within participants. Ordered session-by-session PHQ-9 scores (level 1) were nested within reporting participants (level 2) as outcome. Therefore, MLMs evaluated whether level 1 between session outcome slope (rate of change) was conditional on level 2 (participant) differences, particularly CICS scores. Using established MLM building guidelines (Snijders & Bosker, 2012), model parameters were entered individually, including: Application of random versus fixed intercepts and slopes, baseline PHQ-9 score, session 1 CICS interaction theme scores, treatment group, age, gender, employment status, ethnicity, waiting time, and Index of Multiple Deprivation. Change in $-2 \log$ likelihood was used to identify the model of best fit. Final MLMs included random intercepts and slopes, baseline PHQ-9 score, treatment group, age, gender, and session 1 CICS interaction theme scores as the target covariate. Model residuals were skewed, so confidence intervals were calculated using a 1000 samples bootstrap.

All models controlled for baseline depression score and treatment group (CBT or PCET), with other covariates explored for a model of best fit. Logistic regression was used to predict reliable recovery with last session PHQ-9 reliable recovery achieved (1) or not (0) as outcome (NHS Digital, 2025).

2.7.1. Missing data

Only observed values were used in analyses, and no missing values were imputed, given that the participant inclusion criteria meant sufficient data were available for analysis of included cases. Full information, maximum likelihood methods were applied to MLMs, which allows use of all available data to estimate parameters even in the presence of missing data for some participants. This approach is shown to produce

the best estimates of true effects under conditions of missingness and more accurate assessment of change over repeated assessments than single assessment time-points used in regression analyses (Witkiewitz et al., 2014).

3. Results

3.1. Sample characteristics

Most participants in both conditions were employed, white British women. All participants in both arms were categorized as severely depressed at baseline and reported mean PHQ-9 scores in the moderate-severe range at their first therapy session. At their first therapy session, significantly more severe symptoms of depression were reported on the PHQ-9 by participants in the PCET arm ($M = 19.3$, $SD = 4.7$) compared to CBT ($M = 17.2$, $SD = 5.9$; $z = -1.98$, $p = .047$). There were no other significant demographic or clinical differences between groups at baseline (Table 2).

Clinical outcomes of both therapies were above national averages for reliable recovery: 50% of CBT participants ($n = 20$) and 52.4% of PCET participants ($n = 33$) reporting reliable recovery at their last treatment session (48.3% reliable recovery nationally during trial recruitment; NHS Digital, 2018, Table 2).

Problem and Context Description was the most frequently coded interaction theme in both groups, accounting for a median 60.9% of session 1 utterances in the CBT group and 73.9% in PCET. The least frequently coded interaction theme was Noticing Change or Otherwise, accounting for a median <0.01% of session 1 in both CBT and PCET groups. The interaction theme with the lowest proportion of active engagement from patients was Information Discussion where a median 23.7% of words from first CBT sessions were coded to the theme, but patients only showed active engagement in a median 2.3%. Similarly, 11.6% of session 1 interactions were coded Information Discussion in PCET, but a median of only 2.8% of interactions showed active patient engagement in this theme. Conversely, almost all interactions coded Identifying Protective Factors indicated active patient engagement.

Table 2

Demographics and clinical characteristics of participants receiving CBT and PCET.

	CBT ($n = 40$)	PCET ($n = 63$)
<i>Demographics</i>		
Women	23 (57.5%)	32 (50.7%)
Men	17 (42.5%)	31 (49.2%)
Mean age (SD)	35 (11)	41 (12)
Median IMD Decile (IQR)	6.5 (8)	6 (7)
Ethnicity		
White British	36 (90%)	54 (85.7%)
Other	3 (7.5%)	6 (9.5%)
Unknown	1 (2.5%)	3 (4.8%)
Employment status		
Employed	26 (65%)	41 (65%)
Student/training	5 (12.5%)	0
Unemployed	4 (10%)	8 (12.7%)
Long-term sick or disabled	2 (5%)	6 (9.5%)
Homemaker/not actively seeking work	2 (5%)	4 (6.3%)
Retired	0	4 (6.3%)
Unknown	1 (2.5%)	0
<i>Clinical Characteristics</i>		
Number of sessions attended (SD)	12.6 (5.9)	11.4 (6.2)
Session 1 PHQ-9 (SD)	17.2 (5.9)	19.3 (4.7)
PHQ-9 at last attended session (SD)	9.1 (7.3)	10.7 (6.6)
Number achieving reliable recovery at last session	20 (50%)	33 (52.4%)

Note. Abbreviations: SD, standard deviation; IMD, Index of Multiple Deprivation, IQR, Inter-Quartile Range; CBT, Cognitive Behavioural Therapy; PCET, Person-Centered Experiential Therapy; PHQ-9, Patient Health Questionnaire 9-items.

¹ The second session of therapy was used for the six participants where the first session was not recorded.

There were differences in the patterns of CICS scores between participants in the PCET and CBT groups: In CBT, significantly more of session 1 contained Action Planning and Idea generation interactions with observable patient engagement ($z = 2.15, p = .031$); also in the CBT, a significantly greater proportion of session 1 was rated as structuring and task focus interactions ($z = 3.84, p < .001$). In PCET, a larger proportion of session 1 was coded as Problem or Context Description ($z = 3.54, p < .001$) and Measurement Completion ($z = 3.13, p = .002$; Table 3).

In MLMs controlling for baseline severity, age, gender, and change over time, there were similarities in predictive effects in the CBT and PCET groups (Figs. 2 and 3): Greater patient activation in Action Planning and Idea generation interactions predicted greater improvement in depressive symptoms across therapy sessions in both CBT (PHQ-9, $B = -0.49, SE B = 0.15, p < .001$) and PCET ($B = -0.68, SE B = 0.41, p = .039$), despite accounting for only a median 3.3% and 0.29% of session 1 interactions respectively. Conversely, the larger the proportion of session 1 coded as Measurement Completion, the poorer the outcome across therapy in depressive symptoms for both CBT ($B = 0.27, SE B = 0.07, p < .001$) and PCET ($B = 0.02, SE B = 0.01, p = .003$). Measurement completion accounted for 1.1% and 5.9% of session 1 interactions in CBT and PCET (Table 3).

There were also different patterns of predictive effects in the CBT group and PCET group: Greater patient activation in Identifying Protective Factors interactions at session 1 (accounting for a median 1% of utterances) was associated with greater improvement in depression over therapy sessions in CBT ($B = -0.38, SE B = 0.13, p = .003$), but not PCET ($B = 0.13, SE B = 0.08, p = .042^2$).

In the opposite direction, larger proportions of session 1 coded as Problem or Context Descriptions (CBT $Mdn = 60.9\%$; PCET $Mdn = 73.9\%$) predicted greater improvement in depression for the PCET group ($B = -0.03, SE B = 0.01, p = .002$), but not CBT ($B = 0.03, SE B = 0.03, p > .05$). Larger proportions of session 1 coded as Structuring and Task Focus interactions predicted poorer outcomes in PCET ($B = 0.24, SE B = 0.11, p = .037$), but not CBT ($B = 0.01, SE B = 0.08, p > .05$). This was despite almost double the median percentage of session 1 coded at that theme in CBT ($Mdn = 6.4\%$), compared to PCET ($Mdn = 3.7\%$). Evaluations of Self or Therapy and Information Discussion interactions did not predict outcomes in either group (Figs. 2 and 3; full MLM models are presented in Supplementary Tables 5–18).

Logistic regression analyses did not identify any session 1 CICS themes as predictors of PHQ-9 reliable recovery at the last session in CBT or PCET (Supplementary Tables 3 and 4).

4. Discussion

The present study found that several CICS interaction themes predicted depression improvement, suggesting that patient activation in first session interactions may predict outcomes among severely depressed patients receiving CBT and PCET. Overall, this study suggests that there may be prognostic markers associated with motivation and engagement that are identifiable within what is said between therapists and patients at the first sessions of CBT and PCET.

4.1. Link to previous literature

Commonly utilized CBT manuals identify action-planning and action-orientation as central to the therapeutic approach; practical activities completed outside sessions are a component of almost every part of CBT (Beck, 2020). It is, therefore, understandable that greater improvement in depressive symptoms was associated with greater engagement in action planning in this study. However, more surprising

is the small amount of time where participants were actively engaged in action planning required for this effect (approximately 1% of session 1 interactions) and the early stage at which effects could be identified. Similarly, an emphasis of pre-existing positive aspects of patients' lives, including strengths and protective factors, has been well-developed in CBT manuals (Beck, 2020; Padesky & Mooney, 2012). This may explain why greater patient engagement in identification of protective factors at first sessions predicted outcomes. Yet these interactions again took up only a small amount of session 1 (approximately 1% of session 1 interactions), suggesting that only small changes in practice may be required to integrate this type of benefit early in CBT.

This study is the first to apply the CICS to PCET sessions, which may help to explain the differences from previous study results. Key therapeutic processes in PCET focus on emotional expression and moment-to-moment experiencing, which could explain why greater time characterized as problem and context description was predictive of greater improvement: this could have been time in session 1 where patients' emotional expression was facilitated (Murphy, 2019). Furthermore, an excessive focus on structured assessment processes in session 1 may have restricted moment-to-moment experiencing and emotional expression, which might explain the association between more structuring interactions and poorer outcome at session 1 for PCET.

In both CBT and PCET more time spent completing outcome measures at session 1 predicted poorer outcome. At first this appears surprising, given that robust evidence indicates Routine Outcome Monitoring (ROM) with psychometric symptom or relational assessments, improves treatment effectiveness (Lutz et al., 2021). However, key components of ROM include collecting data, feeding back data, and adapting therapy (Barkham et al., 2023), whereas measurement completion interactions would only constitute data collection – reading items and their responses; other aspects of ROM would fall under different interaction-types. In addition, patients are more supportive of ROM when the information is used in session by therapists, where there is a clear rationale, and when patient engagement is sought (Börjesson & Boström, 2020). These key elements may have been missing in the way measures were completed, as they would have been categorized as different interaction-types.

The current study may aid interpretation of a recent finding using data from the same trial, which found that early improvements on the PHQ-9 (across sessions 1–4) predicted reliable recovery at the end of treatment in PCET, but not in CBT (Arden et al., 2025; Duffy et al., 2022). For example, the beneficial effects of describing problems and their contexts seen in PCET but not CBT may be more weighted towards early sessions, given that the majority of interactions were coded at this theme. Conversely, interactions on the identification of protective factors that predicted improvement in CBT but not PCET accounted for only a small proportion of early discussions. Overall, differences in the two therapies could mean that early session interactions contribute to the earlier versus later change in depressive symptoms observed.

More generally, results from the current study differ from previous research on the CICS's prognostic validity, because two interaction themes that consistently predicted outcomes across measures did not obtain sufficiently reliable coding results: where patients or therapists noticed changes as a result of therapy (Noticing Change or Otherwise) and where an understanding of patient problems was jointly developed (Problem Analysis and Understanding; Malins et al., 2021a; Malins et al., 2021b). Other differences from previous CICS research may relate to the previous focus on second session interactions, where more opportunity to evaluate therapy and notice changes may have occurred, with session structure potentially more orientated towards therapeutic intervention rather than assessment, as would be the case at session 1. Furthermore, two interaction themes predictive of outcomes in the current study had not been defined in previous research: Identifying Protective Factors and Measurement Completion.

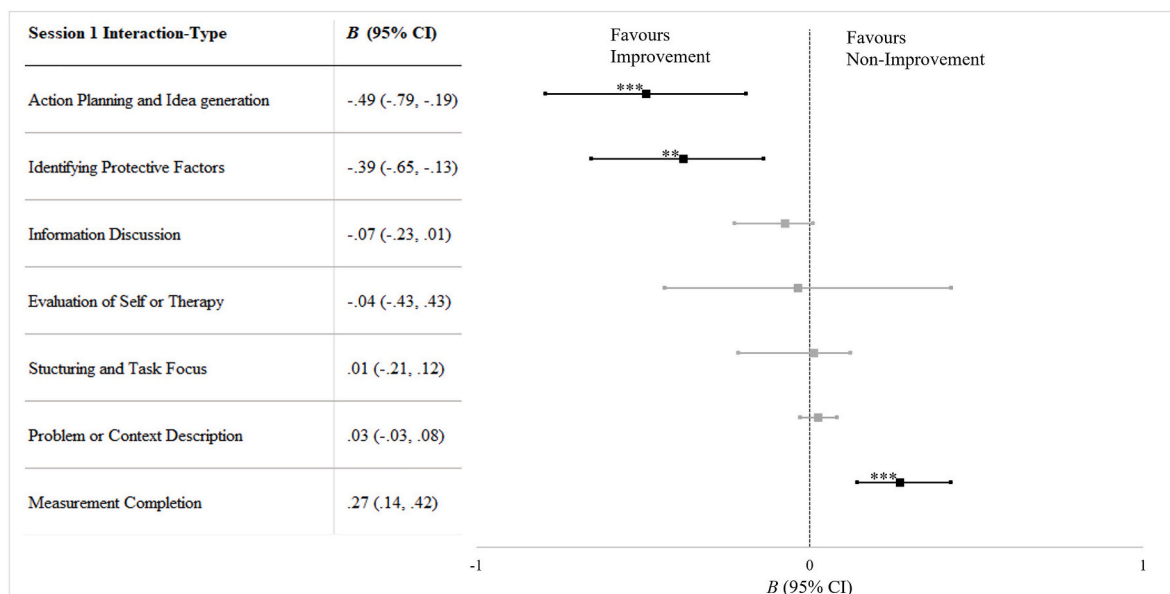
² Treated as a null result, because bootstrapped Confidence Intervals crossed zero (95% CI = $-0.1, 0.3$).

Table 3

Median percentage of session 1 coded at each CICS interaction theme.

CICS Interaction Theme	CBT		PCET	
	Percent Session 1 coded (IQR) ^a	Transformed score (IQR) ^b	Percent Session 1 coded (IQR) ^a	Transformed score (IQR) ^b
Action Planning and Idea generation	3.3 (4.0)***	0.6 (2.2)*	0.29 (1.0)***	<0.01 (0.8)*
Evaluation of Self or Therapy	1.2 (2.0)	0.4 (1.5)	1.00 (1.0)	<0.01 (0.9)
Information Discussion	23.7 (14.0)**	2.3 (4.9)	11.6 (21.0)**	2.8 (6.3)
Identifying Protective Factors	1.0 (2.0)	1.0 (1.8)	0.78 (2.0)	0.8 (2.1)
Measurement Completion	1.1 (3.4)**	- ^c	5.9 (18.2)**	- ^c
Noticing Change or Otherwise	<0.01 (0.0)	<0.01 (0.2)	<0.01 (0.0)	<0.01 (0.0)
Problem Analysis and Understanding	2.9 (4.0)*	1.0 (3.1)	1.4 (4.0)*	0.7 (2.3)
Problem or Context Description	60.9 (14.4)**	- ^c	73.9 (22.6)**	- ^c
Structuring and Task Focus	6.4 (4.0)***	5.9 (4.6)***	3.7 (3.0)***	3.5 (3.0)***

Note. Abbreviation: IQR, Inter-Quartile Range; ^a reports the total percentage of session 1 coded at the given CICS theme overall; ^b reports the score accounting for positive, negative and neutral rated interactions; ^c no transformed score as not rated on a scale just the percentage of session coded.

**Fig. 2.** Session 1 Cognitive Behavioral Therapy Interaction-Types Predicting Session-By-Session PHQ-9 Score.

Note: * $p < .05$ ** $p < .01$. B = unstandardised Beta coefficients from multilevel models with random intercepts and slopes; controlling for baseline severity, change over time, age, and gender.

4.2. Limitations

The inconsistency between MLM and logistic regression results reduces the generalizability of the effects reported, despite MLMs taking better account of missing data than regressions (Witkiewitz et al., 2014). However, this study aimed to generalize in several ways: the focus on PCET alongside CBT, a patient sample with severe depression as the primary problem rather than anxiety, and a primary care therapy setting as opposed to a specialist setting for patients with comorbidities. These differences may offer an alternative explanation for the results observed. Nonetheless, the low inter-rater reliability for some CICS themes rendered them untestable in prediction models, which may suggest that they require clearer definition, or a clearer coder training program may be needed. The differences in familiarity with therapy between the coders (qualified clinical psychologists, versus a researcher and assistant psychologist) may have led to differences in coding interpretation and impacted the reliability of coding. In addition, the fact that the CICS was initially developed for use with CBT may mean that the interaction-theme definitions and rating levels require adaptation to PCET and other therapeutic modalities. For example, the finding that structuring interactions negatively predicted outcomes of PCET, may be because CBT-orientated structuring is misaligned with the PCET approach, but a different type of structuring could still occur that is

defined differently (e.g., process-guiding; Duffy et al., 2024).

The results observed may also be subject to limitations of the CICS coding method: Given that some interaction themes were very rarely identified during the first session (e.g., Noticing Change or Otherwise < 0.01%), achieving inter-rater reliability would be very challenging. Predictive validity is also impacted, because the CICS assesses patients' active engagement in such interactions, but this is not possible if the interactions are entirely absent. Furthermore, overall CICS session scores do not differentiate between shorter segments where patients led the initiation and development of more engaged interactions or longer segments led by therapists with patients contributing (though this is identified on an individual interaction basis). Therefore, the CICS cannot fully discern with certainty the degree to which therapists contribute to patient activation, given its co-constructed nature.

The sample included in this study may be deemed too highly selective, given the smaller proportion of the overall trial participants included. The focus on participants with severe depression may mean that study findings do not generalize to mild or moderate depression. Therefore, future studies specifically focused on mild-to-moderate depression are warranted to address this gap. It is also possible that there may have been systematic differences between CBT and PCET in implementation of session recordings, given the differences between the two groups in the proportion of participants with recordings available.

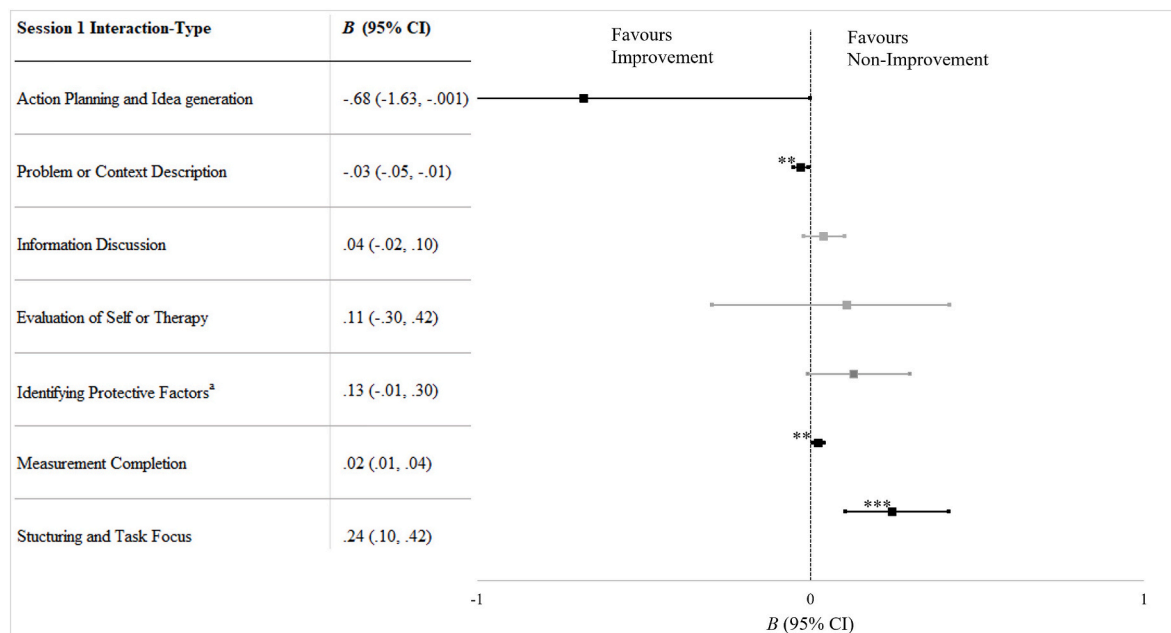


Fig. 3. Session 1 Person-Centered Experiential Therapy Interaction-Types Predicting Session-By-Session PHQ-9 Score.

Note: * $p < .05$ ** $p < .01$ *** $p < .001$. ^atreated as null due to bootstrapped confidence intervals crossing zero. B = unstandardised Beta coefficients from multilevel models with random intercepts and slopes; controlling for baseline severity, change over time, age and gender.

Future studies that include data only from those with more severe and complex problems could include larger proportions of participants. The absence of a formal, direct comparison between CICS scores in CBT and PCET limits the dependability of the differences observed. This was not possible, as the study was not powered for direct comparison, but could be conducted in future research.

4.3. Future research

Future research could more reliably and specifically identify the value of different interaction-types at different time points in therapy through longitudinal ratings, session-by-session across the course of therapy, aiming to predict *next session* outcome – a more proximal measure, alongside the more distal outcome of therapy end, which introduces many confounding variables. This approach would also help to address changes in the proportions of different CICS interaction themes at different points in therapy (e.g. perhaps more noticing change towards the middle-end of therapy). Associations between session-by-session CICS scores and other therapy process measures longitudinally would help to better understand the relation between constructs, the process of change for the patient and the relative predictive validity. Given the differences in predictive validity between interaction-themes in this study, future research should both validate the findings in other samples and test the CICS in other therapeutic modalities.

This study did not focus on the overall, differential clinical effectiveness of CBT or PCET that may be related to CICS scores and related patient activation in early session interactions. Future research could establish if such interaction-types not only predict outcomes, but also differential effectiveness across therapies. Future research should use a larger, purpose-designed sample (either an expanded, coded sample from the parent trial or an independent prospective cohort) to test treatment-by-interaction-type effects directly. Sampling across moderate and severe depression would additionally allow testing of whether these predictive associations vary according to depression severity.

To better facilitate the use of the CICS to support improvement of clinical practice in routine care, future research should continue efforts to automate the reliable rating of therapy sessions using machine learning and natural language processing in explainable ways that

support trust for patients and therapists (Malins et al., 2022). This would allow rapid analysis of large datasets in ways that can both corroborate existing human-coded evidence and can support discovery of new interaction patterns predictive of outcomes.

4.4. Clinical implications

This study supports recent evidence that an initial session, often referred to as an assessment session, may not be best used as solely an information gathering opportunity, but early active engagement in therapeutic processes could support improved outcomes (Aafjes-van Doorn & Sweeney, 2019). Furthermore, the degree to which patients are actively engaged in such therapeutic processes can be a prognostic indicator, using analysis of the language used (Malins et al., 2021a, 2021b).

In both CBT and PCET, engaging patients in action planning, even in very small amounts, at the first session may be beneficial. In CBT, identifying protective factors from the beginning of therapy may improve outcomes and in PCET creating space for patients to describe their problems and context may be beneficial, sacrificing structure to enable this if necessary. Conversely, ROM protocols should be followed to enable therapeutic use of outcome measures, as opposed to administrative completion, which was associated with poorer outcomes across treatments.

The focus on detailed microskills may make the CICS a useful additional tool for deliberate practice methods, enabling therapists to focus on ways of interacting with patients that are prognostic from the beginning of therapy (Miller et al., 2023).

5. Conclusion

This study suggests that patient activation, as measured using the CICS to assess first session interactions, predicts depression across time amongst patients with severe depression receiving CBT and PCET. There are considerable differences in the types of interaction that predict outcome between CBT and PCET, but similarities suggest there may be generally beneficial ways of supporting patient activation across therapeutic modalities from the first session.

CRediT authorship contribution statement

Sam Malins: Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Formal analysis, Conceptualization. **David Saxon:** Methodology, Investigation, Formal analysis, Conceptualization. **Grazziela Figueredo:** Writing – review & editing, Writing – original draft. **Jeremie Clos:** Writing – review & editing, Writing – original draft. **Thomas Trimble:** Writing – review & editing, Writing – original draft. **Kavan Fatehi:** Writing – review & editing, Writing – original draft. **David Waldram:** Writing – review & editing, Writing – original draft, Conceptualization. **Naomi Boycott:** Writing – review & editing, Writing – original draft, Investigation. **Alexandra Frost:** Writing – original draft, Writing – review & editing, Investigation. **Christian Sales:** Writing – review & editing, Writing – original draft, Investigation. **Kate Horton:** Writing – review & editing, Writing – original draft, Conceptualization. **Gillian E. Hardy:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization. **Michael Barkham:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization. **Richard Morriss:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization. **Thomas Schröder:** Writing – review & editing, Writing – original draft, Supervision, Conceptualization. **Nima Moghaddam:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization.

Ethics declaration

Written informed consent to take part in the study and to publish the article has been obtained from all participants or their legal representatives. The privacy rights of participants have been observed.

This study was performed in compliance with relevant laws, regulatory frameworks and guidelines where the research took place. This study was approved by the South Yorkshire Research Ethics Committee. (Approval No. Research Ethics Committee 14/YH/0001)

Data transparency

The data reported in this manuscript draws on a large randomized, non-inferiority controlled trial for which the main outcomes were reported [MS 1]. Subsequently, the data was analysed from the perspective of precision methods using session outcome data [MS 2] and in a test of generalizability of the trial results to the wider clinical organization in which the trial was located [MS 3]. A focus on early sessions was adopted in a report, but only accessed standard PHQ-9 session scores [MS 4]. None of these four published articles utilized analysis of session recordings, as is the case in the current manuscript, which uses therapeutic content to predict outcome, thereby coming closer to linking in-session content to predicted outcomes.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgements

This study was funded through a Health Education England (HEE)/ National Institute for Health and Care Research (NIHR) Clinical Lectureship (Sam Malins, NIHR301292). RM was supported by a NIHR Senior Investigator Award, Nottingham NIHR Biomedical Research Centre, NIHR MindTech HealthTech Research Centre and NIHR Applied Research Collaboration East Midlands. The views expressed are those of the authors and not necessarily those of the NIHR, the NHS or the Department of Health and Social Care. Acknowledgements regarding the PRaCTICED trial itself can be found in the original publication (see Barkham et al., 2021).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.brat.2026.105152>.

Data availability

Data will be made available on request.

References

- Aafjes-van Doorn, K., & Sweeney, K. (2019). The effectiveness of initial therapy contact: A systematic review. *Clinical Psychology Review*, 74, Article 101786. <https://doi.org/10.1016/j.cpr.2019.101786>
- Ardern, K., Baldwin, S. A., Saxon, D., Lorimer, B., Hardy, G. E., & Barkham, M. (2025). Differential effect of early response on outcomes in person-centered experiential therapy and cognitive behavioral therapy for the treatment of adult moderate or severe depression. *Journal of Consulting and Clinical Psychology*, 93(5), 344–356. <https://doi.org/10.1037/ccp0000948>
- Barkham, M., De Jong, K., Delgadillo, J., & Lutz, W. (2023). Routine outcome monitoring (ROM) and feedback: Research review and recommendations. *Psychotherapy Research*, 33(7), 841–855. <https://doi.org/10.1080/10503307.2023.2181114>
- Barkham, M., Saxon, D., Hardy, G. E., Bradburn, M., Galloway, D., Wickramasekera, N., Keetharuth, A. D., Bower, P., King, M., Elliott, R., Gabriel, L., Kellett, S., Shaw, S., Wilkinson, T., Connell, J., Harrison, P., Ardern, K., Bishop-Edwards, L., Ashley, K., Ohlsen, S., Pilling, S., Waller, G., & Brazier, J. E. (2021). Person-centred experiential therapy versus cognitive behavioural therapy delivered in the English Improving Access to Psychological Therapies service for the treatment of moderate or severe depression (PRaCTICED): A pragmatic, randomised, non-inferiority trial. *The Lancet Psychiatry*, 8(6), 487–499. [https://doi.org/10.1016/s2215-0366\(21\)00083-3](https://doi.org/10.1016/s2215-0366(21)00083-3)
- Beck, J. S. (2020). *Cognitive behavior therapy: Basics and beyond* (3rd ed.). Guilford Press.
- Börjesson, S., & Boström, P. K. (2020). “I want to know what it is used for”: Clients’ perspectives on completing a routine outcome measure (ROM) while undergoing psychotherapy. *Psychotherapy Research*, 30(3), 337–347. <https://doi.org/10.1080/10503307.2019.1630780>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>
- Clark, D. M., & Oates, M. (2014). *Improving Access to Psychological Therapies: Measuring improvement and recovery Adult Services. Version 2*. NHS England. <https://www.healthinnovationoxford.org/wp-content/uploads/2015/11/measuring-recovery-2014.pdf>.
- Diamond, G., Wlodek, B., Arthey, S., & Parker, S. (2025). A systematic review of deliberate practice in psychotherapy: Definitions, operationalization, and preliminary outcomes. *Psychotherapy*, 62(2), 113–131. <https://doi.org/10.1037/pst0000569>
- Duffy, K. E. M., Simmonds-Buckley, M., Haake, R., Delgadillo, J., & Barkham, M. (2024). The efficacy of individual humanistic-experiential therapies for the treatment of depression: A systematic review and meta-analysis of randomized controlled trials. *Psychotherapy Research*, 34(3), 323–338. <https://doi.org/10.1080/10503307.2023.2227757>
- Duffy, K. E. M., Simmonds-Buckley, M., Saxon, D., Delgadillo, J., & Barkham, M. (2022). Early response as a prognostic indicator in person-centered experiential therapy for depression. *Journal of Counseling Psychology*, 69(6), 803. <https://doi.org/10.1037/cou0000633>
- Elliott, R., Watson, J. C., Timulak, L., & Sharbanee, J. (2021). Research on humanistic-experiential psychotherapies: Updated review. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (7th ed., pp. 421–467). John Wiley & Sons, Inc.
- Ewbank, M. P., Cummins, R., Tablan, V., Bateup, S., Catarino, A., Martin, A. J., & Blackwell, A. D. (2020). Quantifying the association between psychotherapy content and clinical outcomes using deep learning. *JAMA Psychiatry*, 77(1), 35–43. <https://doi.org/10.1001/jamapsychiatry.2019.2664>
- Ewbank, M. P., Cummins, R., Tablan, V., Catarino, A., Buchholz, S., & Blackwell, A. (2020). Understanding the relationship between patient language and outcomes in internet-enabled cognitive behavioural therapy: A deep learning approach to automatic coding of session transcripts. *Psychotherapy Research*, 1–13. <https://doi.org/10.1080/10503307.2020.1788740>
- Frank, P., Batty, G. D., Pentti, J., Jokela, M., Poole, L., Ervasti, J., Vahtera, J., Lewis, G., Steptoe, A., & Kivimäki, M. (2023). Association between depression and physical conditions requiring hospitalization. *JAMA Psychiatry*, 80(7), 690–699.
- Goldberg, S. B., Babins-Wagner, R., Rousmaniere, T., Berzins, S., Hoyt, W. T., Whipple, J. L., Miller, S. D., & Wampold, B. E. (2016). Creating a climate for therapist improvement: A case study of an agency focused on outcomes and deliberate practice. *Psychotherapy*, 53(3), 367. <https://doi.org/10.1037/pst0000060>
- Goldberg, S. B., Rousmaniere, T., Miller, S. D., Whipple, J., Nielsen, S. L., Hoyt, W. T., & Wampold, B. E. (2016). Do psychotherapists improve with time and experience? A longitudinal analysis of outcomes in a clinical setting. *Journal of Counseling Psychology*, 63(1), 1. <https://doi.org/10.1037/cou0000131>
- Gülüm, İ. V., Dilekter-Aldemir, İ., İkizer, G., & Karanci, A. N. (2025). The temporal dance: How therapist presence and alliance evolve across early sessions of psychotherapy? *Psychology and Psychotherapy: Theory, Research and Practice*, 98, 1017–1032. <https://doi.org/10.1111/papt.70001>

- Hibbard, J., Stockard, J., Mahoney, E. R., & Tusler, M. (2004). Development of the Patient Activation Measure (PAM): Conceptualizing and measuring activation in patients and consumers. *Health Services Research*, 39(4 Pt 1), 1005–1026. <https://doi.org/10.1111/j.1475-6773.2004.00269.x>
- Hill, A. (2011). The competences required to deliver effective counselling for Depression (CfD). *BACP*.
- Hodsoll, J., Strawbridge, R., King, S., Taylor, R. W., Breen, G., Grant, N., Grey, N., Hepgul, N., Hotopf, M., & Kitsune, V. (2024). Predictors of outcome following psychological therapy for depression and anxiety in an urban primary care service: A naturalistic Bayesian prediction modeling approach. *Psychological Medicine*, 54(16), 4503–4517.
- Johns, R. G., Barkham, M., Kellett, S., & Saxon, D. (2018). A systematic review of therapist effects: A critical narrative update and refinement to review. *Clinical Psychology Review*, 67, 78–93. <https://doi.org/10.1016/j.cpr.2018.08.004>
- Kamardina, A., Saunders, R., & 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), 354–374.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Lewis, G. (1994). Assessing psychiatric disorder with a human interviewer or a computer. *Journal of Epidemiology & Community Health*, 48(2), 207–210. <https://doi.org/10.1136/jech.48.2.207>
- Lutz, W., de Jong, K., Rubel, J. A., & Delgadillo, J. (2021). Measuring, predicting, and tracking change in psychotherapy. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (7th ed., pp. 89–133). John Wiley & Sons, Inc.
- Malins, S., Figueredo, G., Jilani, T., Long, Y., Andrews, J., Rawsthorne, M., Manolescu, C., Clos, J., Higton, F., Waldram, D., Hunt, D., Perez Vallejos, E., & Moghaddam, N. (2022). Developing an automated assessment of in-session patient activation for psychological therapy: A co-development approach. *JMIR Medical Informatics*, 10(11), Article e38168. <https://doi.org/10.2196/38168>
- Malins, S., Figueredo, G., Saxon, D., Horton, K., Clos, J., Trimble, T., Fatehi, K., Waldram, D., Higton, F., & Hardy, G. E. (2026). The feasibility and acceptability of automated feedback and deliberate practice in psychological therapies for anxiety and depression. *Digital Health*, 12, Article 20552076251413316. <https://doi.org/10.1177/20552076251413316>
- Malins, S., Moghaddam, N., Morriss, R., Schröder, T., Brown, P., & Boycott, N. (2021a). Predicting outcomes and sudden gains from initial in-session interactions during remote cognitive-behavioural therapy for severe health anxiety. *Clinical Psychology & Psychotherapy*, 28(4), 891–906. <https://doi.org/10.1002/cpp.2543>
- Malins, S., Moghaddam, N., Morriss, R., Schröder, T., Brown, P., & Boycott, N. (2021b). The predictive value of patient, therapist, and in-session ratings of motivational factors early in remote cognitive behavioural therapy for severe health anxiety. *British Journal of Clinical Psychology*, 61(2), 364–384. <https://doi.org/10.1111/bjc.12328>
- Malins, S., Moghaddam, N., Morriss, R., Schröder, T., Brown, P., Boycott, N., & Atha, C. (2020). Patient activation in psychotherapy interactions: Developing and validating the consultation interactions coding scheme. *Journal of Clinical Psychology*, 76, 646–658. <https://doi.org/10.1002/jclp.22910>
- Malins, S., Moghaddam, N., Morriss, R., Schroder, T., Cope, N., & Brown, P. (2026). *Consultation Interaction Coding Scheme 2.0*. Figshare. <https://doi.org/10.6084/m9.figshare.7302386>. Retrieved 12 May from.
- Malins, S., Saxon, D., Figueredo, G., Clos, J., Trimble, T., Fatehi, K., Waldram, D., Boycott, N., Frost, A., Sales, C., Horton, K., Hardy, G. E., Barkham, M., Morriss, R., Schroder, T., & Moghaddam, N. (2026). Analysis code for: Predicting patient outcomes from initial In-Session interactions of cognitive behavioral therapy and experiential therapy for severe depression. Retrieved May 21 from <https://doi.org/10.6084/m9.figshare.32346774>.
- McIntyre, R. S., Alsuwaidan, M., Baune, B. T., Berk, M., Demyttenaere, K., Goldberg, J. F., Gorwood, P., Ho, R., Kasper, S., & Kennedy, S. H. (2023). Treatment-resistant depression: Definition, prevalence, detection, management, and investigational interventions. *World Psychiatry*, 22(3), 394–412.
- Mearns, D., & Thorne, B. (2007). *Person-centred counselling in action* (3rd ed.). Sage.
- Miller, S. D., Chow, D., Malins, S., & Hubble, M. A. (Eds.). (2023). *The field guide to better results: Evidence-based exercises to improve therapeutic effectiveness*. American Psychological Association.
- Miller, S. D., & Hubble, M. A. (2023). Identifying your “what” to practice. In S. D. Miller, D. Chow, S. Malins, & M. A. Hubble (Eds.), *The field guide to better results: Evidence-based exercises to improve therapeutic effectiveness* (pp. 7–24). American Psychological Association. <https://doi.org/10.1037/0000358-002>.
- Moriarty, A. S., Paton, L. W., Snell, K. I., Archer, L., Riley, R. D., Buckman, J. E., Chew Graham, C. A., Gilbody, S., Ali, S., & Pilling, S. (2024). Development and validation of a prognostic model to predict relapse in adults with remitted depression in primary care: Secondary analysis of pooled individual participant data from multiple studies. *BMJ mental health*, 27(1), Article e301226.
- Murphy, D. (2019). *Person-centred experiential counselling for depression: A manual for training and practice* (2nd ed.). Sage Publishing; British Association for Counseling & Psychotherapy.
- National Institute for Health and Care Excellence. (2022). *Depression in adults: Treatment and management*. National Institute for Health and Care Excellence (NICE). <https://www.nice.org.uk/guidance/ng222>.
- NHS Digital. (2018). *Psychological Therapies, annual report on the use of IAPT services - England, 2017-18*. NHS Digital. Retrieved 2 February from <https://digital.nhs.uk/data-and-information/publications/statistical/psychological-therapies-annual-reports-on-the-use-of-iapt-services/annual-report-2017-18>.
- NHS Digital. (2025). *NHS talking Therapies, for anxiety and depression, annual reports, 2024-25*. NHS England. Retrieved 11 May from <https://digital.nhs.uk/data-and-information/publications/statistical/nhs-talking-therapies-for-anxiety-and-depression-annual-reports/2024-25>.
- Padesky, C. A., & Mooney, K. A. (2012). Strengths-based cognitive-behavioural therapy: A four-step model to build resilience. *Clinical Psychology & Psychotherapy*, 19(4), 283–290. <https://doi.org/10.1002/cpp.1795>
- Rhodes, S., Richards, D. A., Ekers, D., McMillan, D., Byford, S., Farrand, P. A., Gilbody, S., Hollon, S. D., Kuyken, W., Martell, C., O'Mahen, H. A., O'Neill, E., Reed, N., Taylor, R. S., Watkins, E. R., & Wright, K. A. (2014). Cost and outcome of behavioural activation versus cognitive behaviour therapy for depression (COBRA): Study protocol for a randomised controlled trial. *Trials*, 15(29). <https://doi.org/10.1186/1745-6215-15-29>
- Rong, J., Wang, X., Cheng, P., Li, D., & Zhao, D. (2025). Global, regional and national burden of depressive disorders and attributable risk factors, from 1990 to 2021: Results from the 2021 global Burden of Disease study. *The British Journal of Psychiatry*, 1–10. <https://doi.org/10.1192/bjp.2024.266>
- Saxon, D., Ashley, K., Bishop-Edwards, L., Connell, J., Harrison, P., Ohlsen, S., Hardy, G. E., Kellett, S., Mukuria, C., Mank, T., Bower, P., Bradburn, M., Brazier, J., Elliott, R., Gabriel, L., King, M., Pilling, S., Shaw, S., Waller, G., & Barkham, M. (2017). A pragmatic randomised controlled trial assessing the non-inferiority of counselling for depression versus cognitive-behaviour therapy for patients in primary care meeting a diagnosis of moderate or severe depression (PRaCTICED): Study protocol for a randomised controlled trial. *Trials*, 18(1), 93. <https://doi.org/10.1186/s13063-017-1834-6>
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Sage.
- Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, 67(6), 894–904. <https://doi.org/10.1037/0022-006x.67.6.894>
- Witkiewitz, K., Falk, D. E., Kranzler, H. R., Litten, R. Z., Hallgren, K. A., O'Malley, S. S., & Anton, R. F., & In conjunction with the Alcohol Clinical Trials Initiative Workgroup. (2014). Methods to analyze treatment effects in the presence of missing data for a continuous heavy drinking outcome measure when participants drop out from treatment in alcohol clinical trials. *Alcoholism: Clinical and Experimental Research*, 38(11), 2826–2834. <https://doi.org/10.1111/acer.12543>