



THE UNIVERSITY *of* EDINBURGH  
School of Health in  
Social Science

## School of Health in Social Science

### Doctorate in Clinical Psychology

|                                                                                             |                                                                                                                              |
|---------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------|
| <b>Name of Course:</b><br>(eg. Clinical Psychology 1)                                       | Research 2                                                                                                                   |
| <b>Type of Assessment:</b><br>(eg. Case Conceptualisation)                                  | Small-Scale Research Project                                                                                                 |
| <b>Title of your Work:</b>                                                                  | Evaluating the Effectiveness of Time-Limited Psychological Therapy under a Service Level Agreement Between NHS Health Boards |
| <b>Exam Number:</b><br>(Please do not include your name or UUN anywhere on this submission) | B253825                                                                                                                      |
| <b>Date of Submission:</b>                                                                  | 22/12/2025                                                                                                                   |
| <b>Word Count:</b>                                                                          | 4037                                                                                                                         |

# Contents

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| Executive Summary .....                                                 | 3  |
| List of Abbreviations.....                                              | 4  |
| Introduction .....                                                      | 5  |
| Service Level Agreement .....                                           | 6  |
| Aims and Objectives .....                                               | 7  |
| Method.....                                                             | 8  |
| Sample.....                                                             | 8  |
| Measures .....                                                          | 8  |
| Analysis .....                                                          | 9  |
| Ethics and Governance .....                                             | 9  |
| Results.....                                                            | 11 |
| Changes in Psychological Distress, Depression and Anxiety Symptoms..... | 12 |
| Reliable Improvement and Post-Treatment Recovery .....                  | 13 |
| Recommendations for Practice .....                                      | 15 |
| Limitations.....                                                        | 16 |
| Improvement Plan .....                                                  | 17 |
| References .....                                                        | 19 |
| Appendix.....                                                           | 22 |

## Executive Summary

**Rationale:** The purpose of this service evaluation was to assess the clinical effectiveness of time-limited cognitive behavioural therapy (CBT) interventions delivered under a Service Level Agreement (SLA) between NHS Orkney and NHS Shetland, which was arranged to reduce service pressure within the NHS Shetland Psychological Therapy (PT) service. The SLA model mandated that a maximum of 12 CBT sessions would be offered per patient, fewer than is typically recommended for individuals with anxiety and more severe depression.

**Aim:** Therefore, this service evaluation aimed to evaluate pre- to post-treatment changes in psychological distress, depression, and anxiety as measured by routinely collected patient outcomes, in order to determine whether the 12-session limit was sufficient for delivery of effective treatment. Findings from this evaluation can then be used to inform future SLA design.

**Method:** The service evaluation used anonymised, routinely collected clinical outcome data from patients who were initially referred under the SLA between May 2024 and May 2025. Following extraction from a digital database, data from a total of 33 adult patients who commenced CBT was analysed.

**Results:** Analyses indicated that patients experienced statistically significant improvements across all outcome measures following CBT delivered under the SLA. Furthermore, more than half of patients demonstrated reliable improvement of depressive symptoms and anxiety, and exactly half demonstrated reliable improvement of overall psychological distress. These findings are discussed, in addition to limitations of the service evaluation, and implications for the service and future SLAs.

## List of Abbreviations

|         |                                                   |
|---------|---------------------------------------------------|
| CMHC    | Common Mental Health Condition                    |
| CBT     | Cognitive Behavioural Therapy                     |
| NICE    | National Institute for Health and Care Excellence |
| SLA     | Service Level Agreement                           |
| PT      | Psychological Therapies                           |
| ROM     | Routine Outcome Monitoring                        |
| CORE-10 | Clinical Outcomes in Routine Evaluation-10        |
| PHQ-9   | Patient Health Questionnaire-9                    |
| GAD-7   | General Anxiety Disorder-7                        |
| SPSS    | Statistical Package for Social Sciences           |

## Introduction

It is estimated that one in five (20.2%) adults in England are impacted by a common mental health condition (CMHC), the most prevalent of which are depression and anxiety disorders, with rates of 3.8% and 7.5% respectively (NHS England, 2025). According to the NHS, overall CMHC rates have also increased in prevalence among 16 to 64 year olds over a 20 year period, from 15.5% in 1993 to 22.6% in 2023/4, and these increasing rates place additional pressure on mental health services to deliver interventions effectively. To aid in the improvement of the delivery and quality of evidence-based psychological interventions for CMHCs, the Scottish Government (2023) published a specification which recommended several strategies to be implemented by mental health services including the routine measurement of outcomes in treatment, and the review of treatment efficacy of psychological interventions.

Cognitive behavioural therapy (CBT) is a widely recommended evidence-based psychological intervention for treating CMHCs like depression and anxiety by the National Institute for Health and Care Excellence (NICE, 2011, 2022), and is routinely administered in stepped-care psychological treatment in NHS services. CBT has a strong evidence base supporting its effectiveness across various presentations, including complex and comorbid difficulties (Bauer et al., 2012; Cuijpers et al., 2016). NICE guidelines advise that individuals who are referred for high-intensity psychological interventions for anxiety should receive between 12 to 15 weekly sessions of CBT, although treatment length should be adapted depending on individual patient requirements (NICE, 2011). For depression, treatment length may depend on symptom severity, with 8 sessions of individual CBT recommended for less severe depression, or 16 regular sessions advised for more severe depression (NICE, 2022). However, implementing these recommendations may be challenging in services that encounter challenges related to geographical constraints or workforce shortages.

Health Boards that cater to remote and rural communities encounter unique difficulties in providing steady access to psychological treatments like CBT, especially when there are workforce limitations: research has shown that worldwide, mental health professional shortages are greatest in rural areas compared to more urban areas (Hoeft et al., 2018), and that access to mental health care services is lower for individuals in rural areas in the UK (Inghels et al., 2025). Issues such as reduced staffing capacity may limit the availability or duration of therapeutic interventions that can be offered to patients in these areas, which has been found to negatively impact patient safety (O'Dowd, 2024). Consequently, these challenges pose a real risk to rural

patients, have prompted impacted services to sometimes adopt alternative solutions, including cross-service agreements.

### **Service Level Agreement**

In response to reduced staffing capacity within NHS Shetland Psychological Therapies (PT) service, a Service Level Agreement (SLA) was established with NHS Orkney PT service. Under this agreement, NHS Orkney clinicians remotely delivered up to 12 sessions of CBT to referred NHS Shetland patients. The primary aim of this agreement was to improve patient access to evidence-based psychological therapy during a period of service pressure, consequently minimising waiting times and disruption to care.

The SLA provided NHS Orkney PT service input for new patients on the NHS Shetland secondary care Community Mental Health Team PT waiting list, as well as for patients who were open but had not completed treatment following previous psychologist input, in line with NHS Orkney PT service referral criteria. Although a six-month period between 01 May 2024 and 31 October 2024 was initially agreed upon, the SLA was subsequently extended into 2025 due to its impact on reducing waiting list pressures and breaches. While this arrangement proved effective in addressing service pressures, the SLA model mandated that a maximum of 12 CBT sessions would be offered per patient, fewer than is typically recommended for both individuals with anxiety and more severe depression under NICE guidance (NICE, 2011, 2022). This reduction compared to recommended treatment duration highlights the importance of evaluating whether such these service adaptations are still capable of delivering meaningful clinical change within routine practice.

Routine outcome monitoring (ROM) is an established practice of NHS PT services that can aid in the evaluation of clinical change and contribute to service improvement. There is also some evidence that the use of ROM possibly improves patient outcomes by a small effect size (McAleavey et al., 2024). Common measures include the CORE-10 (Clinical Outcomes in Routine Evaluation-10; Barkham et al., 2013), PHQ-9 (Patient Health Questionnaire-9; Kroenke et al., 2001), and GAD-7 (General Anxiety Disorder-7; Spitzer et al., 2006), which are validated tools for assessing psychological distress, depression, and anxiety, respectively, and each are used by the NHS Orkney PT service for ROM. Due to these measures' pre-existing routine use within the PT service, and their sensitivity to changes in psychological distress and for diagnosis of anxiety and depression (Barkham et al., 2013; Villarreal-Zegarra et al., 2023), these were chosen for analysis of clinical change under the SLA.

## **Aims and Objectives**

This report aims to assess the clinical effectiveness of time-limited CBT interventions delivered under the SLA between NHS Orkney and NHS Shetland by examining pre- to post-treatment changes in psychological distress, depression, and anxiety as measured by routinely collected CORE-10, PHQ-9, and GAD-7 patient outcomes. The following four questions were identified:

1. Are there significant pre–post changes in psychological distress following CBT delivered under the SLA, as measured by the CORE-10?
2. Are there significant pre–post changes in depressive and anxiety symptoms following CBT delivered under the SLA, as measured by the PHQ-9 and GAD-7?
3. What proportion of patients receiving CBT under the SLA demonstrate reliable improvement on outcome measures?
4. What proportion of patients receiving CBT under the SLA demonstrate post-treatment recovery on outcome measures?

These findings will be used to provide service improvement guidance in relation to future implementation of SLAs for PT service provision to other Health Boards.

## Method

### Sample

The present research is a retrospective pre-post service evaluation using anonymised, routinely collected clinical outcome data from patients who were initially referred under the SLA between May 2024 and May 2025. Data was extracted from a digital dataset completed for each patient referred under the SLA to the NHS Orkney PT service for remotely-delivered therapy by NHS Orkney therapists. A total of 68 cases were referred to the NHS Orkney PT service under the SLA, although of these, only 35 had a recorded number of total sessions, and 33 of these had at least one baseline outcome measure recorded.

All patients met the NHS Orkney PT service referral criteria, which included adults ( $\geq 18$  years) whose mental health difficulties were likely to respond to a time limited period of high intensity psychological intervention, and whose goals could be met through psychological intervention. Presenting problems may have included, but were not exclusive to: depressive disorders, anxiety disorders (including social anxiety, health anxiety, panic disorder, and phobias), OCD, trauma/complex trauma/PTSD, or low self-esteem. Exclusion criteria included individuals with presentations outside the service remit such as active psychosis, substance misuse, personality disorders, or relationship problems.

### Measures

During treatment, ROM was carried out with patients to assess changes in presentation. This included use of the CORE-10 (measuring general psychological distress on a scale of 0-40), the PHQ-9 (depressive symptom severity on a scale of 0-27), and the GAD-7 (anxiety severity on a scale of 0-21). The CORE-10 defines scores of 0-5 as healthy, 6-10 as low level (non-clinical), 11-14 as mild, 15-19 as moderate, 20-24 as moderate-to-severe, and 25+ as severe psychological distress (Barkham et al., 2013). The PHQ-9 defines scores of 0-4 as minimal, 5-9 as mild (non-clinical), 10-14 as moderate, 15-19 as moderately severe, and 20-27 as severe (Kroenke et al., 2001), whilst the GAD-7 defines scores of 0-4 as minimal, 5-9 as mild, 10-14 as moderate, and 15-21 as severe (Spitzer et al., 2006).

In addition to these variables, start and end dates of therapy, total sessions, and patient gender were also extracted. Due to limitations of the digital database used for ROM, reasons for discharge were not made available, nor was patient age.

## **Analysis**

All analyses were conducted in the Statistical Package for Social Sciences (SPSS; IBM Corp, 2025). To analyse pre-post changes in psychological distress, depression and anxiety symptoms, t-tests were selected. Change scores were calculated as post - pre, with negative values denoting improvement. Statistical tests were two-tailed with  $\alpha = .05$ . All significance tests were conducted with a significance level of 5%, and missing data was handled using pairwise exclusion, allowing for all available data to be utilised.

Reliable improvement was calculated according to the reliable change index for each measure. This was defined as a decrease of 6 or more points for the CORE-10 (Barkham et al., 2013), a decrease of 6 or more points for the PHQ-9, and a decrease of 4 or more points for the GAD-7 (Porter et al., 2024).

Post-treatment recovery was calculated as post-treatment scores below standard clinical thresholds for each measure. These thresholds are defined as scores  $\leq 10$  for the CORE-10, scores  $\leq 9$  for the PHQ-9, and scores  $\leq 7$  for the GAD-7. A McNemar's test was planned to assess the change in the proportion of sub-clinical CORE-10 scores from pre- to post-treatment. However, the test could not be carried out for the cases with paired data, as all paired cases were above the clinical threshold at pre-treatment, and consequently there were no cases below clinical thresholds at baseline to compare post-treatment recovery against.

Although total sessions was extracted, analyses examining the relationship between total number of sessions and outcomes were not conducted, as in clinical practice the session count reflects patient need and clinical judgement, rather than an independent treatment “dose”, meaning it would be difficult to interpret any observed effects. Similarly, although patient gender was extracted, it was not included in outcome analyses as this evaluation was not designed to examine differential treatment effects by demographic characteristics, and subgroup analyses were not sufficiently powered due to small sample size.

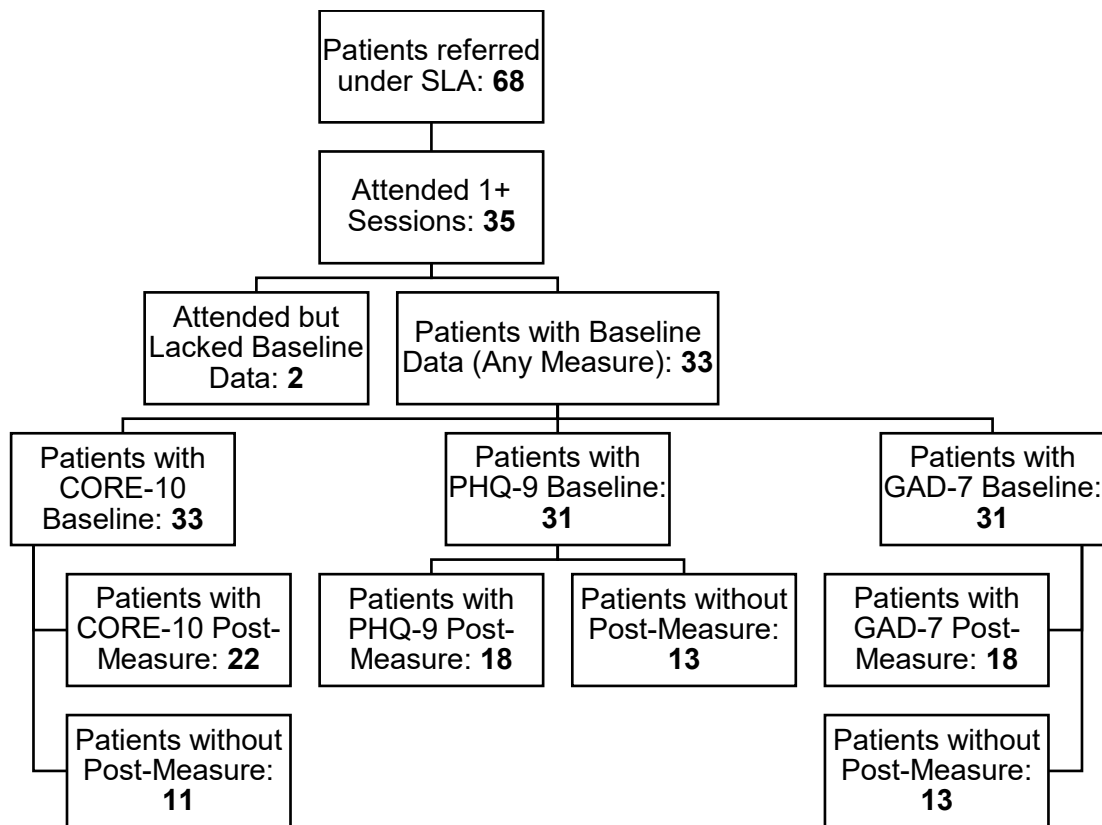
## **Ethics and Governance**

This present evaluation was conducted as a service evaluation using anonymised, routinely collected clinical data, and did not involve collection of new patient data, and therefore did not require NHS Research Ethics Committee review. The SLA between NHS Orkney and NHS Shetland included an agreement for authorisation of audit/evaluation activity and data sharing for proper conduct of audit work within each organisation. Caldicott approval was granted by the NHS

Orkney Clinical Governance Team, and following this, ethical approval was granted from the University of Edinburgh, School of Health in Social Science Research Ethics Committee (see appendix). All data was anonymised prior to transfer, and stored electronically on a secure university network only accessible by the primary researcher. Files will be retained and deleted according to University of Edinburgh data-retention policy. Ethical principles of confidentiality, integrity, and data protection were maintained throughout, and all data handling complied with GDPR and local NHS information-governance policies. Findings will be shared with both Health Boards and used for service development purposes.

## Results

Outcome measures were collected from a total of 33 adult patients who commenced cognitive behavioural therapy (CBT) under the SLA between NHS Orkney and NHS Shetland. All patients met NHS Orkney PT referral criteria and were offered treatment remotely. As this evaluation relied on routinely collected clinical data, not all patients had complete pre- and post-treatment outcome measures across all scales. Patient flow and data inclusion across outcome measures are summarised in Figure 1 below.



**Figure 1.** CONSORT-Style Flowchart of Patient Progression and Data Inclusion

Baseline outcome data was available for 33 patients on the CORE-10 and for 31 patients on the PHQ-9 and GAD-7. Paired pre- and post-treatment data was available for CORE-10 ( $n = 22$ ), PHQ-9 ( $n = 18$ ), and GAD-7 ( $n = 18$ ), which were used for pre-post outcome comparisons.

Across the evaluation period, a total of 174 CBT sessions delivered under the SLA were registered on the digital database. The mean number of sessions attended per patient was 4.97 ( $SD = 4.23$ ), with attendance ranging from 1 to 14 sessions. Although the SLA specified a maximum of 12

sessions per patient, this was exceeded in a single instance, with 14 sessions offered to one patient, which was retained in analyses.

### Changes in Psychological Distress, Depression and Anxiety Symptoms

At baseline, patients demonstrated clinically significant levels of psychological distress, depression, and anxiety. The mean CORE-10 score at assessment was 22.00 (SD = 6.13), indicating an average of “moderate-to-severe” psychological distress. Mean baseline PHQ-9 and GAD-7 scores were 17.50 (SD = 4.23) and 14.44 (SD = 3.53) respectively, indicating an average of at least “moderate” severity for both depressive symptoms and anxiety. These findings suggest that the sample represents a clinically complex population.

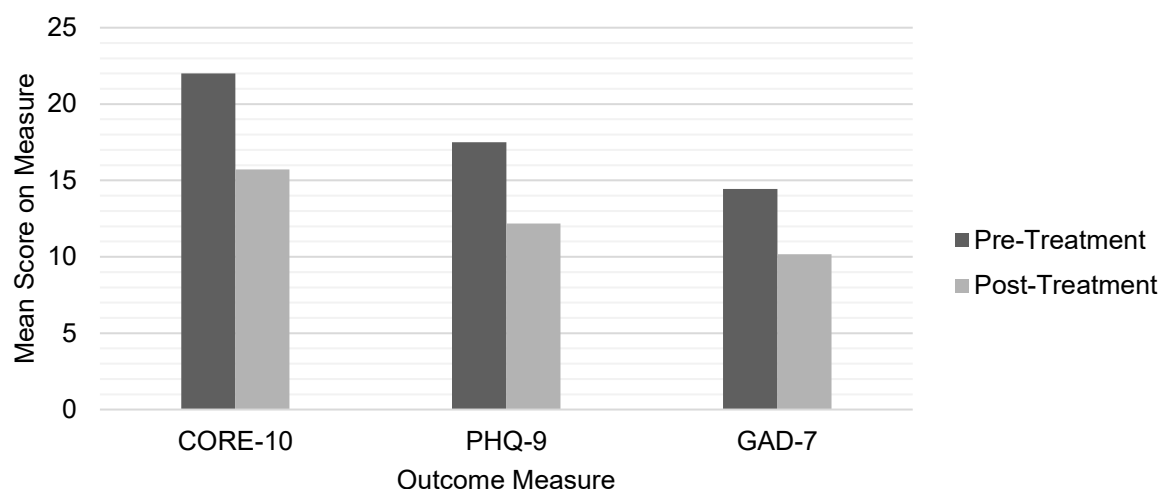
Change scores (pre-post) for each outcome measure were assessed for normality using the Shapiro–Wilk test. Results indicated that normality assumptions were not violated for any measure: CORE-10 change ( $p = .229$ ), PHQ-9 change ( $p = .090$ ), and GAD-7 change ( $p = .115$ ). Therefore, parametric paired-samples  $t$ -tests were considered appropriate for all pre-post outcome analyses.

**Table 1.** Pre- and Post-Treatment Mean Scores, Mean Change, and Inferential Statistics

| Measure        | N  | Pre-Treatment<br>Mean (SD) | Post-Treatment<br>Mean (SD) | Mean<br>Change | $t$ (df)      | $p$   | Cohen's<br>$d$ |
|----------------|----|----------------------------|-----------------------------|----------------|---------------|-------|----------------|
| <b>CORE-10</b> | 22 | 22.00 (6.13)               | 15.73 (7.60)                | -6.27          | 4.713<br>(21) | <.001 | 1.01           |
| <b>PHQ-9</b>   | 18 | 17.50 (4.23)               | 12.17 (6.45)                | -5.33          | 4.331<br>(17) | <.001 | 1.02           |
| <b>GAD-7</b>   | 18 | 14.44 (3.54)               | 10.17 (5.46)                | -4.28          | 3.851<br>(17) | 0.001 | 0.91           |

Pre- and post-treatment outcome statistics are presented in Table 1 above. Paired-samples  $t$ -tests demonstrated statistically significant improvements across all outcome measures. For the CORE-10, there was a significant reduction in scores from pre- to post-treatment,  $t(21) = 4.71$ ,  $p < .001$ , with a large effect size (Cohen's  $d = 1.01$ ), indicating a reduction in overall psychological distress. Similarly, PHQ-9 scores decreased significantly following treatment,  $t(17) = 4.33$ ,  $p < .001$ , with a large effect size ( $d = 1.02$ ), suggesting improvement in depressive symptoms. For

GAD-7, a significant reduction in anxiety symptoms was also observed,  $t(17) = 3.85$ ,  $p = .001$ , with a large effect size ( $d = 0.91$ ). These findings indicate that patients experienced statistically significant improvements following CBT delivered under the SLA. This improvement across all measures is visualised in Figure 2 below.

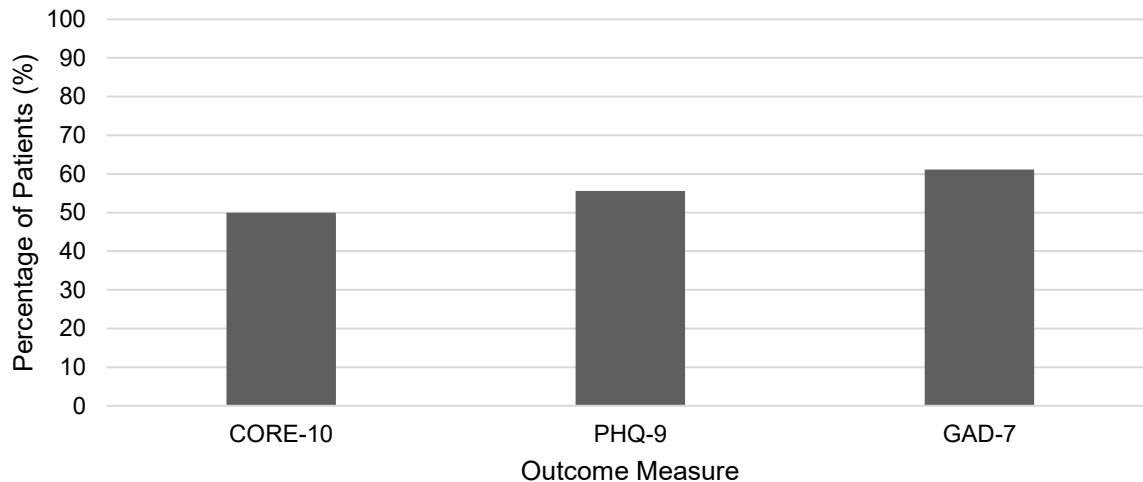


**Figure 2.** Mean scores for each measure at pre- and post- treatment.

Attrition bias was explored descriptively by comparing baseline outcome scores for all patients who completed pre-treatment measures with those included in paired pre-post analyses. For the CORE-10, the baseline mean score for all patients completing the measure ( $N = 33$ ) was 21.00 ( $SD = 6.65$ ), compared to 22.00 ( $SD = 6.13$ ) in the paired analysis sample ( $N = 22$ ). Baseline PHQ-9 scores were similarly comparable between the full cohort ( $N = 31$ ;  $M = 17.32$ ,  $SD = 5.22$ ) and the paired sample ( $N = 18$ ;  $M = 17.50$ ,  $SD = 4.23$ ), as were GAD-7 scores between the full cohort ( $N = 31$ ;  $M = 14.55$ ,  $SD = 4.49$ ) and paired sample ( $N = 18$ ;  $M = 14.44$ ,  $SD = 3.54$ ). Overall, baseline severity appeared comparable between patients included in outcome analyses and those lost to follow-up.

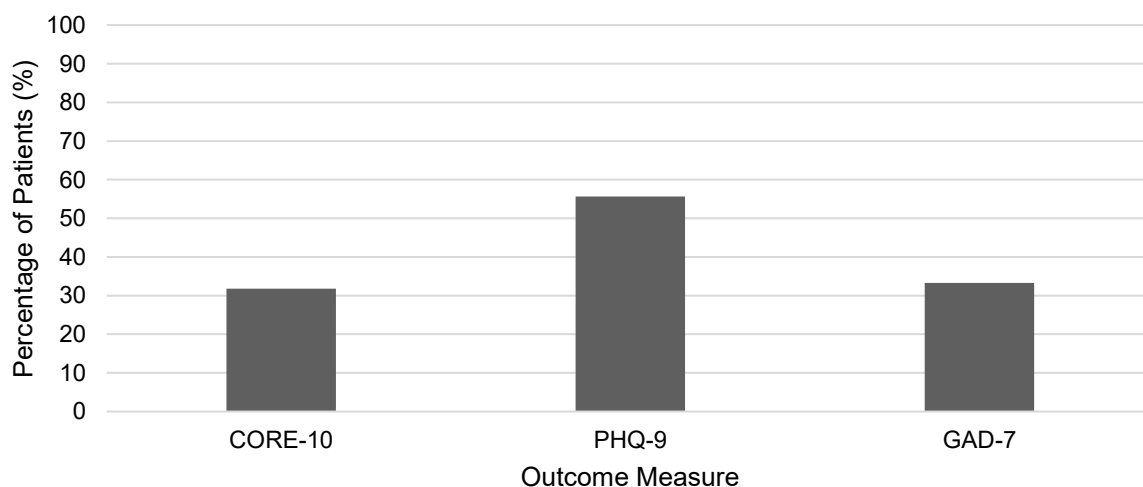
### Reliable Improvement and Post-Treatment Recovery

Rates of reliable improvement were calculated using established reliable change thresholds for each outcome measure. On the CORE-10, 50.0% (11/22) of patients achieved reliable improvement. Reliable improvement was observed in 55.6% (10/18) of patients on the PHQ-9 and 61.1% (11/18) of patients on the GAD-7 at post-treatment. These rates of reliable improvement are visualised in Figure 3 below.



**Figure 3.** *Proportion of patients achieving reliable improvement post-treatment.*

Post-treatment scores were also examined in relation to established clinical cut-off thresholds. For all cases with paired scores, no patients were below the clinical threshold at pre-treatment. On the CORE-10, 31.8% (7/22) of patients scored below the clinical threshold at post-treatment. Recovery rates were higher for patients with depressive symptoms, with 55.6% (10/18) of patients scoring below the PHQ-9 clinical cut-off at post-treatment, and 33.3% (6/18) of patients with clinically significant anxiety scores achieving recovery on the GAD-7 at post-treatment. These findings indicate that whilst a substantial proportion of patients achieved non-clinical symptom levels at discharge, others continued to experience symptoms of psychological distress, anxiety and depression. These recovery rates are visualised in Figure 4 below.



**Figure 4.** *Proportion of patients below clinical cut-off post-treatment*

## Recommendations for Practice

The purpose of this service evaluation was to explore the effectiveness of CBT interventions delivered under the SLA between NHS Orkney and NHS Shetland, which was introduced due to staff pressures in the NHS Shetland PT service. Results indicated that there were significant improvements across all outcome measures, with large effect sizes observed for each. In addition to this, more than half of patients demonstrated reliable improvement on the PHQ-9 and GAD-7, and exactly half demonstrated reliable improvement on the CORE-10. Notably, the rate of reliable improvement amongst patients under the SLA aligns with expected figures, as approximately 50-to-66% of patients are found to improve following psychological intervention according to Castonguay et al. (2021). Furthermore, between 31.8% and 55.6% patients demonstrated post-treatment scores that fell below established clinical thresholds, with rates varying between each measure. Overall, these findings suggest that treatment delivered under the SLA was successful in producing meaningful outcomes.

NICE guidelines recommend CBT as a first-line intervention for CMHC like depression and anxiety, with lengthier (12+ session) interventions advised for more complex difficulties, or individuals with comorbid conditions (NICE, 2011, 2022). Given that the maximum number of sessions allocated under the SLA (12) fell below those recommended by best practice treatment guidance for CMHCs like severe depression and anxiety, the present evaluation offers important evidence regarding the impact of these constraints and capped session count on patient outcomes. Despite these limits, meaningful improvements were observed, supporting the feasibility of future SLA arrangements with similar sessions allocations. These findings also reinforce existing research on the efficacy of reduced-length treatments: brief CBT lasting three to six sessions has previously been linked to meaningful improvements in depression symptoms (Cully et al., 2024), and similarly, research has found brief 8-session CBT to be feasible and beneficial for the treatment of anxiety disorders (Crawley et al., 2013). Notably, it has also been found that briefer interventions can result in lower attrition rates for treatment (Öst & Ollendick, 2017), highlighting further potential benefits of such treatment models.

It is important to note that although reduced-length treatment can still lead to meaningful clinical outcomes, as evidenced by the present evaluation, they may not always be appropriate for individuals with more complex presentations. Furthermore, research has found that larger doses of therapy are able to help in maintaining gains and reducing relapse rates (Kagan et al., 2021). It is possible that increased availability of sessions for those requiring additional input would have

resulted in greater improvements and a larger proportion of patients demonstrating sub-clinical scores at post-treatment. This underscores the need for continued flexibility in mental health services and the importance of allowing clinicians to make patient-centred decisions around treatment length. Importantly, despite the allocated 12-session cap, there was one instance wherein treatment was extended to a total of 14 sessions. Information on the factors behind this decision was not made available, but it is plausible that this arose from routine clinical considerations such as risk management or consolidation of therapeutic gains. Additionally, given that the average number of sessions delivered to each patient was considerably fewer than was maximally allocated under the SLA, this instance is unlikely to have had a meaningful impact on overall clinician capacity, which the maximum allocation sought to protect. As such, this might support the value of formally allowing clinicians greater flexibility in decisions around session length under future SLAs, particularly as overly restrictive allocations may dissuade most clinicians from offering additional sessions that could benefit a patient, even if capacity exists.

## **Limitations**

One key limitation of this evaluation was the incomplete outcome data for several patients. Although pairwise exclusion was used in analyses in order to maximise the use of the available data, missing post-treatment scores reduced paired sample sizes substantially for all measures, which may have resulted in an impact on observed results. It is possible that patients who lacked post-treatment outcomes may have experienced poorer outcomes, subsequently resulting in overestimation of treatment effectiveness in analyses; alternatively, some of these patients may have improved over the course of therapy, but the lack of their post-treatment measures could have attenuated the observed effect of treatment and limited accurate interpretation of outcomes. However, as noted, mean baseline scores for all patients who completed pre-treatment measures did not differ substantially from the scores of those included in pre-post analyses, so the impact of attrition bias (which would suppose that those without post-treatment scores may have differed systematically from those with post-treatment scores) is likely limited, although it cannot fully be excluded.

It is crucial to highlight that the digital database for patient outcomes lacked comprehensive information about reason for discharge or completion of treatment, which further obfuscates the reason for absence of post-treatment scores. In many cases, missing post-treatment scores may be due to patient disengagement, early discharge that was agreed between the patient and clinician (for example, if a patient's presentation had changed since referral and therefore psychological intervention was no longer appropriate), or a clinician may have not collected or

uploaded further outcome measures during the remainder of treatment. However, under all of these instances, the patient would still simply be marked as “discharged” with no further information available. In many cases, this information might be found on separate patient systems, such as those that hold clinical continuation forms, but this qualitative data would need to be extracted manually from a variety of sources and lengthen the time required to carry out the evaluation substantially. Ideally, alterations to the digital database for patient outcomes, perhaps integrating a drop-down menu with preset categories of discharge reason, would be beneficial for streamlining future evaluations of this kind.

A further consequence of missing post-treatment outcome data was the relatively small sample size, which was insufficient for meaningful subgroup analyses (such as those of gender differences), and would have likely increased the risk of both false positive and false negative findings (Burke et al., 2015). This small sample size may also limit the generalisability of this evaluation to some degree, although to an extent this is a reflection of the limited number of patients from NHS Shetland who accessed treatment under the SLA.

Additionally, the lack of follow-up outcomes limits the interpretation of durability of observed treatment effects; the use of patient questionnaires administered after a set period following discharge could provide useful information regarding the maintenance of gains made by patients in treatment, and future evaluations may wish to incorporate such measures if feasible. Furthermore, whilst patient outcomes provide useful information on the clinical effectiveness of treatment delivered under the SLA, future service evaluations may also wish to explore the impact of the SLA on NHS Orkney waiting times and overall clinical outcomes, particularly as one of the primary aims of the SLA was to mitigate service pressure within the NHS.

## **Improvement Plan**

In context of these findings and limitations, an improvement plan is proposed below:

1. The findings of this service evaluation should be disseminated to both the NHS Orkney and NHS Shetland PT services, in order to raise awareness of the clinical effectiveness of treatment delivered under the SLA and inform future SLA design. This should also serve to highlight the importance and utility of routine outcome measures in service governance and evaluation, in addition to their clinical use.
2. Completion of outcome measures (CORE-10, PHQ-9, and GAD-7) should be embedded into routine discharge practice, in order to reduce missing data and improve the validity of

future evaluations. This may be supported by the introduction of a structured discharge checklist to encourage consistency amongst clinician outcome recording procedure.

3. Data quality and recording standards should be improved by agreeing minimum data standards within future SLA documentation. If feasible, the addition of standardised fields for discharge reason should be implemented into the digital database for outcoming recording, in order to support interpretation of outcomes in future evaluations and reduce reliance of manual data extraction.
4. Under future SLA agreements, service evaluation and outcome monitoring expectations should be embedded explicitly into future SLA agreements. This may include planning review points for data completeness during SLAs.
5. Session caps should be revisited under future SLAs. It is recommended that limited clinical flexibility around session limits should be retained where justified, as a fixed 12 session limit may unnecessarily restrict clinician decision-making when capacity exists to offer additional sessions in accordance with patient need.
6. Evaluation of clinical effectiveness should be repeated under future SLAs, with follow-up measures incorporated in order to assess the maintenance of therapeutic gains. In addition to this, future evaluations may wish to expand their scope to include service-level indicators, such as impact on overall NHS Orkney PT service waiting times.

## References

- Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., Mellor-Clark, J., Richards, D., Unsworth, G., & Evans, C. (2013). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling & Psychotherapy Research*, 13(1), 3–13. <https://doi.org/10.1080/14733145.2012.729069>
- Bauer, I., Wilansky-Traynor, P., & Rector, N. A. (2012). Cognitive-Behavioral Therapy for Anxiety Disorders with Comorbid Depression: A Review. *International Journal of Cognitive Therapy*, 5(2), 118–156. <https://doi.org/10.1521/ijct.2012.5.2.118>
- Burke, J. F., Sussman, J. B., Kent, D. M., & Hayward, R. A. (2015). Three simple rules to ensure reasonably credible subgroup analyses. *BMJ*, 351, h5651. <https://doi.org/10.1136/bmj.h5651>
- Castonguay, L. G., Barkham, M., Youn, S. J., & Page, A. C. (2021). Practice-based evidence—Findings from routine clinical settings. In *Bergin and Garfield's handbook of psychotherapy and behavior change: 50th anniversary edition, 7th ed* (pp. 191–222). John Wiley & Sons, Inc.
- Crawley, S. A., Kendall, P. C., Benjamin, C. L., Brodman, D. M., Wei, C., Beidas, R. S., Podell, J. L., & Mauro, C. (2013). Brief Cognitive-Behavioral Therapy for Anxious Youth: Feasibility and Initial Outcomes. *Cognitive and Behavioral Practice*, 20(2), 123–133. <https://doi.org/10.1016/j.cbpra.2012.07.003>
- Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M., & Huibers, M. J. H. (2016). How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry*, 15(3), 245–258. <https://doi.org/10.1002/wps.20346>
- Cully, J. A., Hundt, N. E., Fletcher, T., Sansgiry, S., Zeno, D., Kauth, M. R., Kunik, M. E., & Sorocco, K. (2024). Brief Cognitive-Behavioral Therapy for Depression in Community Clinics: A Hybrid Effectiveness-Implementation Trial. *Psychiatric Services (Washington, D.C.)*, 75(3), 237–245. <https://doi.org/10.1176/appi.ps.20220582>
- Hoefl, T. J., Fortney, J. C., Patel, V., & Unützer, J. (2018). Task-Sharing Approaches to Improve Mental Health Care in Rural and Other Low-Resource Settings: A Systematic Review. *The Journal of Rural Health*, 34(1), 48–62. <https://doi.org/10.1111/jrh.12229>
- IBM Corp. (2025). *IBM SPSS Statistics for Windows* (Version 31.0) [Computer software].

- Inghels, M., Nelson, D., Kane, R., Gussy, M., & Deaney, C. (2025). Impact of rural-urban residence and deprivation on care pathways for depression disorders among adults in the UK. *The Journal of Rural Health*, 41(3), e70055. <https://doi.org/10.1111/jrh.70055>
- Kagan, E. R., Frank, H. E., Knepley, M. J., & Kendall, P. C. (2021). Beyond 16 Sessions: Extending Manualized Treatment of Anxious Youth. *Journal of Child and Family Studies*, 30(2), 493–501. <https://doi.org/10.1007/s10826-020-01872-5>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine*, 16(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- McAleavey, A. A., de Jong, K., Nissen-Lie, H. A., Boswell, J. F., Moltu, C., & Lutz, W. (2024). Routine Outcome Monitoring and Clinical Feedback in Psychotherapy: Recent Advances and Future Directions. *Administration and Policy in Mental Health*, 51(3), 291–305. <https://doi.org/10.1007/s10488-024-01351-9>
- NHS England. (2025, November 27). *Chapter 1: Common mental health conditions*. NHS England Digital. <https://digital.nhs.uk/data-and-information/publications/statistical/adult-psychiatric-morbidity-survey/survey-of-mental-health-and-wellbeing-england-2023-24/common-mental-health-conditions>
- NICE. (2011, January 26). *Recommendations | Generalised anxiety disorder and panic disorder in adults: Management | Guidance | NICE*. NICE. <https://www.nice.org.uk/guidance/cg113/chapter/Recommendations>
- NICE. (2022, June 29). *Recommendations | Depression in adults: Treatment and management | Guidance | NICE*. NICE. <https://www.nice.org.uk/guidance/ng222/chapter/Recommendations>
- O'Dowd, A. (2024). Staff shortages have worsened safety of mental health patients by preventing therapeutic care. *BMJ (Clinical Research Ed.)*, 387, q2361. <https://doi.org/10.1136/bmj.q2361>
- Öst, L.-G., & Ollendick, T. H. (2017). Brief, intensive and concentrated cognitive behavioral treatments for anxiety disorders in children: A systematic review and meta-analysis. *Behaviour Research and Therapy*, 97, 134–145. <https://doi.org/10.1016/j.brat.2017.07.008>
- Porter, A., Franklin, M., De Vocht, F., d'Apice, K., Curtin, E., Albers, P., & Kidger, J. (2024). Estimating the effectiveness of an enhanced 'Improving Access to Psychological Therapies' (IAPT) service addressing the wider determinants of mental health: A real-world evaluation. *BMJ Open*, 14(1), e077220. <https://doi.org/10.1136/bmjopen-2023-077220>

- Scottish Government. (2023, September 29). *Psychological therapies and interventions specification*. <https://www.gov.scot/publications/psychological-therapies-interventions-specification/>
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Villarreal-Zegarra, D., Barrera-Begazo, J., Otazú-Alfaro, S., Mayo-Puchoc, N., Bazo-Alvarez, J. C., & Huarcaya-Victoria, J. (2023). Sensitivity and specificity of the Patient Health Questionnaire (PHQ-9, PHQ-8, PHQ-2) and General Anxiety Disorder scale (GAD-7, GAD-2) for depression and anxiety diagnosis: A cross-sectional study in a Peruvian hospital population. *BMJ Open*, 13(9), e076193. <https://doi.org/10.1136/bmjopen-2023-076193>

# Appendix



SCHOOL of HEALTH IN SOCIAL SCIENCE

The University of Edinburgh  
Medical School  
Doorway 6, Teviot Place  
Edinburgh EH8 9AG

Email [ethics.hiss@ed.ac.uk](mailto:ethics.hiss@ed.ac.uk)

01 December 2025

Dear Edward Pulleyn,

**Reference:** 25-26CLPS032

**Project Title:** Evaluating the Effectiveness of Time-Limited Psychological Therapy in Complex Cases: An Outcome Study of a Service Level Agreement Between NHS Orkney and NHS Shetland

Thank you for submitting the above research project for review by the School of Health in Social Science Research Ethics Committee (REC). I can confirm that the submission has been independently reviewed and has received a favourable opinion on 27 November 2025.

The standard conditions of this are:

- I. Conduct the project strictly in accordance with the proposal that you have submitted and that has been granted ethics favourable opinion, including any amendments made to the proposal required by the REC.
- II. Advise the REC (by email to [ethics.hiss@ed.ac.uk](mailto:ethics.hiss@ed.ac.uk)) of any complaints or other issues in relation to the project, which may warrant review of the favourable opinion granted to the project.
- III. Make submission for approval of amendments to the project before implementing such changes.
- IV. Advise in writing if the project has been discontinued.

The School's Research Ethics Policy and further information and resources are available on the School's website.

Best of luck with your project.

Yours sincerely,

Sam Mills

Research Support Administrator  
School of Health in Social Science