



Nottingham Trent
University
Psychology

Analysis of Project-ITQ Closed Cases

Authors: Mark Andrews

Version: e9aeba9

Date: 30 November, 2025

Abstract

This report presents analyses of treatment outcomes for 335 military veterans with closed cases in Project-ITQ. Clients were predominantly male (74.7%), with a median age of 49.5 years. Approximately 54% had experienced combat, 42% were employed, and 43% were taking antidepressants. Treatment outcomes were assessed using GAD-7, PHQ-9, PCL-5, PRN-14, and the International Trauma Questionnaire (ITQ). Substantial improvements were observed across all measures, with large effect sizes (Cohen's $d = 1.0-1.14$) and reliable improvement rates of 57–69%. Recovery rates on combined GAD-7/PHQ-9 criteria reached 40%. Analysis of ITQ diagnostic trajectories revealed differential recovery patterns: 90% of clients initially diagnosed with PTSD achieved “No PTSD” status at treatment completion, compared to only 53% of those with Complex PTSD. Logistic regression controlling for initial symptom severity identified living situation and referral source as modest predictors of recovery, with clients living with parents showing higher recovery rates, potentially reflecting social support benefits. ANCOVA analyses demonstrated that initial ITQ diagnosis did not predict differential treatment response on GAD-7, PHQ-9, or PCL-5 once baseline severity was controlled; final scores were well predicted by initial scores alone (slope $\approx 0.4-0.5$). These findings indicate that while Complex PTSD clients have lower absolute recovery rates, this is attributable to greater initial severity rather than reduced treatment responsiveness. Treatment trajectories are equivalent across diagnostic categories when accounting for baseline symptom levels.

1 Client details

There were 335 clients with closed cases. For some or all clients, we have information about their gender, age, ethnicity, employment status, living arrangements, medication use, and how they were referred to their treatment. For their employment status, living arrangements, medication use, and referral status, their responses, if any, were in free form. This made it difficult or impossible to have a limited and meaningful number of values for each variable and as such, it was necessary to simplify or aggregate their responses concerning these variables, as we explain below.

1.1 Gender

Of all the 332 clients, 248 (74.7%) were listed as male, 66 (19.9%) were listed as female. For 16 (4.8%) clients, their gender was not stated.

1.2 Age

Of the 332 clients, 330 provided their age. Their median age was 49.5, with 95% of cases being between 27 and 72.77. Their mean age was 48.99 with a standard deviation was 12.36. The minimum age was 22 and the maximum age was 98.

1.3 Ethnicity

Overall, 285 clients (85.8%) stated their ethnicity as *White British*. Of those remaining, 21 did not state their ethnicity, and 8 stated their ethnicity as *Other White*. The total breakdown by ethnicity category is as follows:

Ethnicity	n
White British	285 (85.8%)
Not known/stated	21 (6.3%)
Other White	8 (2.4%)
White Irish	4 (1.2%)
Caribbean	3 (0.9%)
Other Asian	2 (0.6%)
Other Mixed	2 (0.6%)
White and Black African	2 (0.6%)
African	1 (0.3%)
Any Other	1 (0.3%)
Other Black	1 (0.3%)
White and Asian	1 (0.3%)
White and Black Caribbean	1 (0.3%)

1.4 Employment status

If clients stated their employment status as either “employed”, “self-employed”, or “student”, they were classified as “employed”. If they stated their employment status as “long-term sick” or “short-term sick”, their employment status was classified as “sick”. If they stated they were seeking work, their status was classified as “seeking”, and if they stated they were either retired or not seeking work, they were classified as “not seeking”. Any other responses concerning employment were classified as “other”. The breakdown of these categories is as follows:

Employment status	n
employed	139 (41.5%)
sick	78 (23.3%)
not seeking	47 (14%)
other	46 (13.7%)
seeking	25 (7.5%)

1.5 Military service

At screening, 320 clients provided a free text response to the question *Where did you serve?* Their responses were simplified into three categories: “combat”, meaning that they experienced active military duty in some military conflict theatre; “no combat”, meaning that they did not serve in a military conflict theatre; “other”, meaning any other response. Clients who mentioned in their free response terms like *Northern Ireland, NI, Afghan(istan), Gulf, Iraq, Falklands, Yugoslavia, Bosnia, or Kosovo* were classified as “combat”. If they explicitly stated *no active tours*, they were classified as “no combat”. All other responses, were classified as “other”. The breakdown of these categories is as follows:

Military service	n
combat	172 (53.8%)
other	116 (36.2%)
no combat	32 (10%)

1.6 Living arrangements

Many clients stated that their living arrangements were one of the following: “with partner”, “house/flat share”, “with parents”, “alone”, “homeless”. The first three of these responses were relabelled as “partner”, “share”, “parents”, respectively, and the response “Wife and children” was also classified as “partner”. All other responses, each of which had only one response across clients, were classified as “other”. The breakdown is as follows:

Living arrangements	n
partner	146 (43.6%)
alone	110 (32.8%)
other	43 (12.8%)
parents	25 (7.5%)
homeless	7 (2.1%)
shared	4 (1.2%)

1.7 Medication use

In response to a question about medication use, 142 (42.8%) clients indicated they were taking anti-depressants, 1 indicated they were taking anti-psychotics, and 1 indicated they were taking anxiolytics or hypnotics. All others indicated they were taking no medication, or “other” medication, or did not respond. Because of the absence of much detail about medication use, we treated this as just a binary variable that indicated whether the clients was taking anti-depressants or not. As mentioned, 142 (42.8%) stated they were taking anti-depressants and so 190 (57.2%) were recorded as not taking anti-depressants.

1.8 Referral status

In response to a question about how they were referred for treatment, most either explicitly indicated that they self-referred or else they did not respond at all. However, those cases where there was no response, were treated as self-referrals. The remaining clients listed many different sources for their referral, such as *Care After Combat*, *RBLI*, *Red Poppy Factory*, but each of these other values usually had a very small number of respondents, often just one each. To simplify, therefore, we treat referral status as a binary variable indicating whether a client self-referred or not. Overall, 198 (59.1%) self-referred, and so 137 (40.9%) did not.

2 Session details

In total, 321 (96.7%) clients did one or more (post-screening, prior to follow-up) therapy sessions with one of 73 different therapists. The median number of sessions done by these clients was 6, and 90% of clients did between 1 and 12 sessions. Of these 321 clients, 55 (17.1%) did not complete the therapy programme as planned, i.e., they dropped-out before completion.

3 Analysis of improvements in outcome measures

For the GAD-7, PHQ-9, and PCL-5 outcome variables, we summarized and analysed improvement while in therapy, i.e. from initial assessment to final assessment excluding assessment done at follow-up times, using an IAPT style approach. For each variable, we calculated the number of clients who can be included in the analysis, the number of initial clinical “cases”, the number of recovered cases, the number of reliably recovered cases, the number of clients who reliably improved. We also analysed improvement using a paired-samples t-test and calculated the Cohen’s d ¹ and Hedges’ g score. These results are summarized in Table 1 and are explained in more detail in the following subsections.

In addition, following standard IAPT guidelines (see NHS England, 2019, p. 53), we report the number of clients who are initially defined as cases by either GAD-7 or PHQ-9 definitions and who are then recovered cases on both measures at the end of therapy. In particular, we used the following definitions:

- A client is defined as a combined GAD-7 and PHQ-9 case if they are defined as a case on GAD-7 or PHQ-9 or both.
- A client is defined as a recovered combined GAD-7 and PHQ-9 case if at the end of therapy, they are not a case on GAD-7 and also not a case on PHQ-9, i.e. below the caseness threshold on both measures.
- A client is defined as reliably improved if at the end of therapy their GAD-7 score or their PHQ-9 or both have improved by more than the improvement threshold, and their GAD-7 nor their PHQ-9 scores have not reliably deteriorated².
- A client is defined as reliably recovered if they are both recovered and have reliably improved.

In addition to the standard outcome measures described above, we also analysed the International Trauma Questionnaire [ITQ; Cloitre et al. (2018)]. The ITQ is a self-report measure designed to assess symptoms of PTSD and Complex PTSD (CPTSD) according to the ICD-11 diagnostic criteria. Unlike traditional single-score measures, the ITQ provides two primary subscale scores: a PTSD score (assessing re-experiencing, avoidance, and hyperarousal) and a Disturbances in Self-Organization (DSO) score (assessing affective dysregulation, negative self-concept, and disturbances in relationships). For this analysis, we examined changes in both subscale scores from initial to final measurements using paired t-tests and effect size calculations.

3.1 GAD-7

- There were 288 clients who provided initial measurements of GAD-7. Of these, 252 (87.5%) clients were measured at least one more time and so can be included in this analysis.
- Of the clients for whom we have at least two measures, 223 (88.5%) were cases, defined as having a score of 8 or above, at the time of their first measurement.

¹In the context of paired initial and final scores, Cohen’s d effect size is defined, see Cohen (1988), as follows:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2 + s_2^2 - 2rs_1s_2}} \times \sqrt{2(1 - r)},$$

where $\bar{x}_1$ is the mean of the initial scores, $\bar{x}_2$ is the mean of the final scores, s_1 is the standard deviation of the initial scores, s_2 is the standard deviation of the final scores, and r is the Pearson correlation coefficient between the initial and the final scores.

²From the IAPT manual, “Patients are considered reliably improved if their scores for depression and/or anxiety have reduced by a reliable amount and neither measure has shown a reliable increase.”

- Of those clients who were initial cases, 102 (45.7%) had recovered, i.e. were no longer cases, by the time of their final measurement.
- Of these recovered cases, 99 (44.4%) had reliably recovered in that they had both recovered and their score had improved by at least 4 points.
- Of all the clients with two or more GAD-7 measures, 155 (61.5%) reliably improved in that their score improved by at least 4 points from the initial to the final measurement, regardless of whether their initial or final measurement was above or below the caseness threshold.
- The average (and standard deviation) improvement from first to final measure was 6.22 (5.91).
- A paired t-test comparing the initial and final scores shows that there was a highly significant improvement on average: $t(251) = 16.71, p < 0.001$.
- The Cohen's d effect size for the average improvement is $d = 1.14$. The Hedges' g correction of this Cohen's d effect size is $g = 1.13$.

3.2 PHQ-9

- There were 272 clients who provided initial measurements of PHQ-9. Of these, 235 (86.4%) clients were measured at least one more time and so can be included in this analysis.
- Of the clients for whom we have at least two measures, 192 (81.7%) were cases, defined as having a score of 10 or above, at the time of their first measurement.
- Of those clients who were initial cases, 87 (45.3%) had recovered, i.e. were no longer cases, by the time of their final measurement.
- Of these recovered cases, 84 (43.8%) had reliably recovered in that they had both recovered and their score had improved by at least 6 points.
- Of all the clients with two or more PHQ-9 measures, 135 (57.4%) reliably improved in that their score improved by at least 6 points from the initial to the final measurement, regardless of whether their initial or final measurement was above or below the caseness threshold.
- The average (and standard deviation) improvement from first to final measure was 7.07 (6.81).
- A paired t-test comparing the initial and final scores shows that there was a highly significant improvement on average: $t(234) = 15.93, p < 0.001$.
- The Cohen's d effect size for the average improvement is $d = 1.03$. The Hedges' g correction of this Cohen's d effect size is $g = 1.02$.

3.3 PCL-5

- There were 253 clients who provided initial measurements of PCL-5. Of these, 181 (71.5%) clients were measured at least one more time and so can be included in this analysis.
- Of the clients for whom we have at least two measures, 149 (82.3%) were cases, defined as having a score of 32 or above, at the time of their first measurement.
- Of those clients who were initial cases, 79 (53%) had recovered, i.e. were no longer cases, by the time of their final measurement.
- Of these recovered cases, 76 (51%) had reliably recovered in that they had both recovered and their score had improved by at least 10 points.
- Of all the clients with two or more PCL-5 measures, 125 (69.1%) reliably improved in that their score improved by at least 10 points from the initial to the final measurement, regardless of whether their initial or final measurement was above or below the caseness threshold.
- The average (and standard deviation) improvement from first to final measure was 20.04 (18.07).
- A paired t-test comparing the initial and final scores shows that there was a highly significant improvement on average: $t(180) = 14.93, p < 0.001$.
- The Cohen's d effect size for the average improvement is $d = 1.08$. The Hedges' g correction of this Cohen's d effect size is $g = 1.07$.

Table 1: Summary of IAPT style outcome analyses for GAD-7, PHQ-9, and PCL-5.

	GAD-7	PHQ-9	PCL-5
Initial n	288	272	253
Included n	252 (87.5%)	235 (86.4%)	181 (71.5%)
Initial cases	223 (88.5%)	192 (81.7%)	149 (82.3%)
Recovered cases	102 (45.7%)	87 (45.3%)	79 (53%)
Reliable recovery	99 (44.4%)	84 (43.8%)	76 (51%)
Reliable improvement	155 (61.5%)	135 (57.4%)	125 (69.1%)
Average improvement	6.22 (5.91)	7.07 (6.81)	20.04 (18.07)
Paired t-test	$t(251) = 16.71, p < 0.001$	$t(234) = 15.93, p < 0.001$	$t(180) = 14.93, p < 0.001$
Cohen's d	1.14	1.03	1.08
Hedges' g	1.13	1.02	1.07

3.4 GAD-7 & PHQ-9 combined

- There were 272 clients who provided initial measurements of both GAD-7 and PHQ-9. Of these, 235 (86.4%) clients were measured at least one more time on both and so can be included in this analysis.
- Of the clients for whom we have at least two measures on both, 212 (90.2%) were cases, defined as being a case on either GAD-7 or PHQ-9 or on both, at the time of their initial measurements.
- Of those clients who were initial cases, 86 (40.6%) had recovered, i.e. they were not cases on both GAD-7 and PHQ-9, by the time of their final measurement.
- Of these recovered cases, 84 (39.6%) had reliably recovered in that they had recovered by the end of therapy, i.e. were not cases on both GAD-7 and PHQ-9, and either their GAD-7 or PHQ-9 scores or both had reliably improved and neither scores had reliably deteriorated.

These results are summarized in Table 2.

Table 2: Summary of combined GAD-7 and PHQ-9 outcomes.

	GAD-7 & PHQ-9
Initial n	272
Included n	235 (86.4%)
Initial cases	212 (90.2%)
Recovered cases	86 (40.6%)
Reliable recovery	84 (39.6%)

3.5 PRN-14

- There were 255 clients who provided initial measurements of PRN-14. Of these, 180 (70.6%) clients were measured at least one more time and so can be included in this analysis.
- The average (and standard deviation) improvement³ from first to final measure was 23.23 (29.62).
- A paired t-test comparing the initial and final scores shows that there was a highly significant improvement on average: $t(179) = 10.52, p < 0.001$.
- The Cohen's d effect size for the average improvement is $d = 0.75$.

³Unlike with GAD-7, PHQ-9, PCL-5, CORE-10, higher scores in the PRN-14 measure are better outcomes. To simplify comparisons with the other measures, improvements in PRN-14 are calculated by subtracting the initial score from the final score, while in the other measures, improvement scores are calculated from subtracting the final score from the initial score.

- The Hedges' g correction of this Cohen's d effect size is $g = 0.747$.

3.6 ITQ scores

- There were 263 clients who provided initial measurements of ITQ. Of these, 175 (66.5%) clients were measured at least one more time and so can be included in this analysis.
- The ITQ does not provide one single overall score but provides two major subscores: *PTSD score* and *DSO score*.
- For the PTSD scores, the average (and standard deviation) improvement from first to final measure was 5.65 (5.3).
- A paired t-test comparing the initial and final scores shows that there was a highly significant improvement on average: $t(174) = 14.1, p < 0.001$.
- The Cohen's d effect size for the average improvement is $d = 1.00$. The Hedges' g correction of this Cohen's d effect size is $g = 0.99$.
- On the other hand, for the DSO scores, the average (and standard deviation) improvement from first to final measure was 6.12 (5.76).
- A paired t-test comparing the initial and final scores shows that there was a highly significant improvement on average: $t(174) = 14.06, p < 0.001$.
- The Cohen's d effect size for the average improvement is $d = 1.00$. The Hedges' g correction of this Cohen's d effect size is $g = 1.00$.

These results are summarized in Table 3.

Table 3: Summary of ITQ score changes from initial to final assessment.

	DSO	PTSD
Initial n	263	263
Included n	175 (66.5%)	175 (66.5%)
Average improvement	6.12 (5.76)	5.65 (5.3)
Paired t-test	$t(174) = 14.06, p < 0.001$	$t(174) = 14.1, p < 0.001$
Cohen's d	1.00	1.00
Hedges' g	1.00	0.99

3.7 ITQ diagnosis

The ITQ provides dimensional scores for PTSD and DSO (Disturbances in Self-Organization) symptoms, which can be used to derive categorical diagnoses of No PTSD, PTSD, or Complex PTSD based on ICD-11 diagnostic criteria. The diagnostic algorithm considers both symptom endorsement patterns and functional impairment.

To examine how ITQ diagnoses change from initial to final assessment, we analyzed data from 175 clients who had at least two ITQ assessments during therapy (excluding follow-up sessions). Table 4 shows the percentage of clients in each initial diagnostic category who transitioned to each final diagnostic category.

Table 4: Percentage of clients transitioning between ITQ diagnostic categories from initial to final assessment.

diagnosis	n	No PTSD	PTSD	Complex PTSD
No PTSD	50	96.0%	0.0%	4.0%
PTSD	21	90.5%	4.8%	4.8%
Complex PTSD	104	52.9%	10.6%	36.5%

The results show substantial differences in outcomes across initial diagnostic categories. Of the 50 clients initially diagnosed with No PTSD, 48 (96.0%) remained at No PTSD at final assessment. Of the 21 clients initially diagnosed with PTSD, 19 (90.5%) recovered to No PTSD. In contrast, of the 104 clients initially diagnosed with Complex PTSD, only 55 (52.9%) recovered to No PTSD, while 38 (36.5%) remained Complex PTSD.

To formally test differences in recovery rates, we conducted a logistic regression analysis with recovery (achieving No PTSD status at final assessment) as the outcome and initial diagnosis as the predictor. The estimated probability of recovery for clients initially diagnosed with No PTSD was 0.96 (95% CI: 0.85–0.99), for PTSD was 0.90 (95% CI: 0.69–0.98), and for Complex PTSD was 0.53 (95% CI: 0.43–0.62). These results are illustrated in Figure 1.

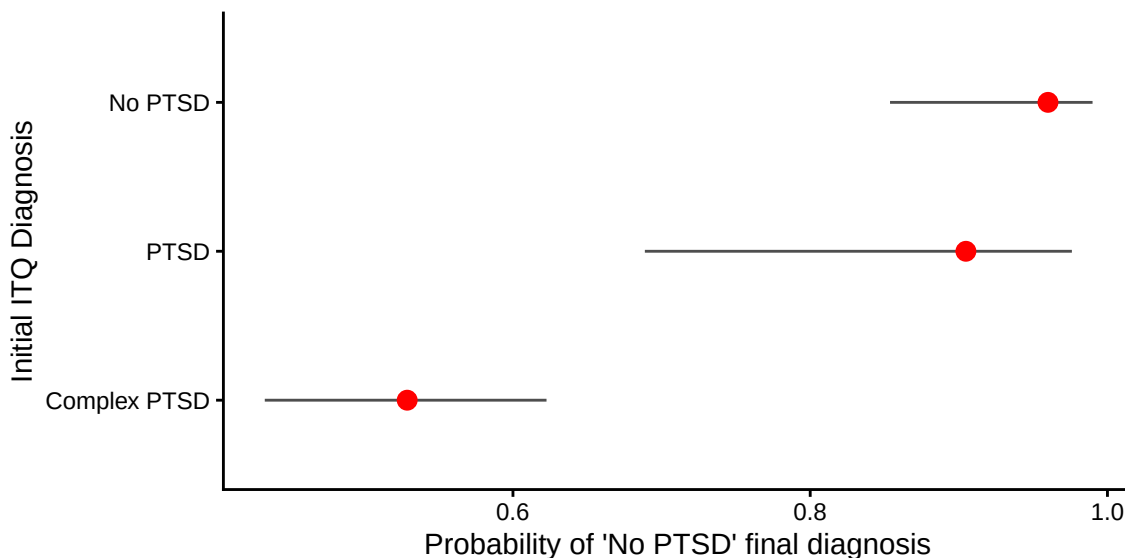


Figure 1: Estimated probability of recovery (achieving “No PTSD” diagnosis at final assessment) by initial ITQ diagnosis. The intervals represent 95% confidence intervals.

3.7.1 Predictors of recovery

Given the substantial difference in recovery rates between clients with PTSD versus Complex PTSD diagnoses, we investigated whether demographic and personal characteristics could help identify factors associated with recovery. Recovery was defined as achieving a “No PTSD” diagnosis at final assessment. We used logistic regression to model recovery as a function of available client characteristics: age, gender, living situation, self-referral status, employment status, use of antidepressant medication, and military service type. Critically, we controlled for initial severity by including each client’s initial PTSD and DSO symptom scores from the ITQ, as clients with more severe initial presentations may have different recovery trajectories regardless of demographic factors.

We excluded three clients in the “shared” or “homeless” living categories due to insufficient data for reliable estimation. The full model included all available predictors, and we used stepwise selection (based on AIC) to identify the most parsimonious model.

Stepwise selection retained four predictors: initial PTSD score, initial DSO score, living situation, and self-referral status. The analysis was based on 160 clients with complete data on all variables. Both initial severity measures were significant predictors: higher initial PTSD scores ($\beta_{\text{initial_ptsd}} = -0.2, p = 0.001$) and higher initial DSO scores ($\beta_{\text{initial_dso}} = -0.16, p = 0.005$) were associated with

lower odds of recovery. This confirms that the lower recovery rates among Complex PTSD clients are largely attributable to their greater initial symptom severity.

Beyond initial severity, two demographic predictors showed meaningful associations with recovery. Living situation was a significant predictor, with clients living with parents showing higher recovery rates compared to those living alone. Self-referral status showed a marginal effect, with self-referred clients having somewhat lower recovery rates.

To aid interpretation, we computed estimated marginal probabilities of recovery for each level of these predictors, averaging over the other covariates using proportional weighting to reflect the sample composition (Table 5 and Table 6).

Table 5: Estimated marginal probabilities of recovery by living situation.

Living situation	Probability	SE	95% CI lower	95% CI upper
Alone	0.727	0.074	0.563	0.846
Other	0.922	0.062	0.686	0.985
With parents	0.970	0.034	0.766	0.997
With partner	0.820	0.053	0.692	0.902

Clients living with parents had the highest estimated probability of recovery (0.97), followed by those in “other” living arrangements (0.92), with a partner (0.82), and alone (0.73). This pattern suggests that social support may play a protective role in recovery.

Table 6: Estimated marginal probabilities of recovery by referral type.

Referral type	Probability	SE	95% CI lower	95% CI upper
Not self-referred	0.891	0.044	0.772	0.952
Self-referred	0.783	0.052	0.664	0.868

Clients who were not self-referred had a somewhat higher estimated probability of recovery (0.89) compared to self-referred clients (0.78), though this difference was only marginally significant.

In summary, after controlling for initial symptom severity, most demographic and personal characteristics (age, gender, employment status, antidepressant use, and military service type) did not significantly predict recovery. The primary determinant of recovery was initial severity: clients with lower initial PTSD and DSO scores were more likely to achieve a “No PTSD” diagnosis at treatment completion. Living situation showed a modest association, with those living with parents having better outcomes, potentially reflecting the benefits of social support during treatment.

4 Individual differences in treatment outcomes by ITQ diagnosis

To investigate whether clients’ initial ITQ diagnosis (No PTSD, PTSD, or Complex PTSD) predicts differential treatment response, we examined final scores on GAD-7, PHQ-9, and PCL-5. Because clients with different ITQ diagnoses typically present with different baseline severity on these measures, and final scores naturally depend on initial scores, we used ANCOVA to compare outcomes while controlling for initial scores. This allows us to determine whether ITQ diagnosis predicts treatment trajectory independent of baseline severity.

Figure 2 shows scatterplots of final scores against initial scores for each measure, with points coloured by initial ITQ diagnosis and separate regression lines fitted for each diagnostic group.

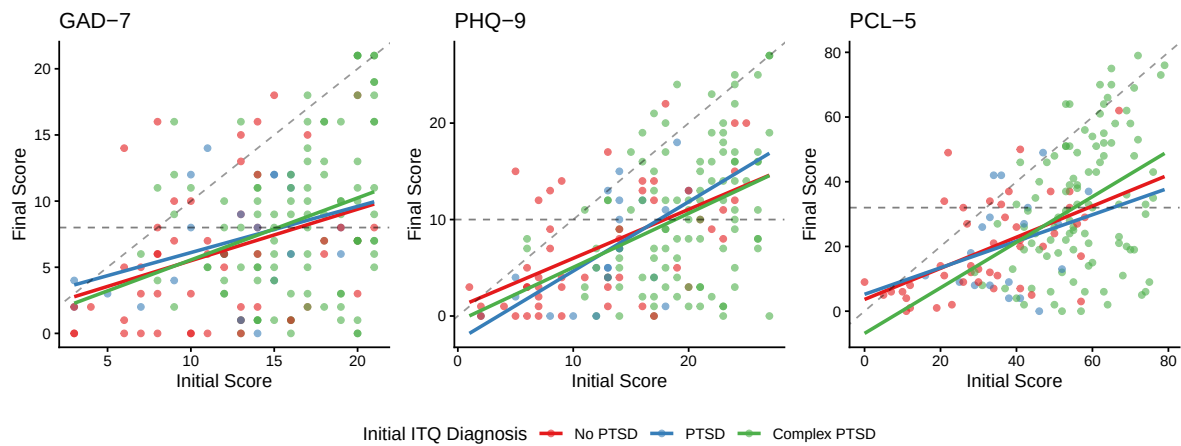


Figure 2: Final scores as a function of initial scores for GAD-7, PHQ-9, and PCL-5, coloured by initial ITQ diagnosis. Dashed diagonal lines represent no change (identity line), dashed horizontal lines indicate clinical caseness thresholds. Regression lines are fitted separately for each ITQ diagnostic group.

The plots reveal a consistent pattern across all three measures. Nearly all points fall below the identity line, indicating that most clients improved during therapy. Points fill the lower triangle bounded by the identity line and the x-axis, reflecting a floor effect: clients with lower initial scores have less room to improve than those with higher initial scores. Importantly, the regression lines for all three ITQ diagnostic groups are closely overlapping, suggesting similar intercepts and slopes regardless of initial diagnosis.

To formally test the effect of initial ITQ diagnosis, we conducted ANCOVA analyses for each measure, with the results summarised in Table 7, which clearly indicates that final scores on each measure do not vary with initial ITQ diagnosis once scores on the initial measures are known.

Table 7: ANCOVA results testing whether initial ITQ diagnosis predicts final scores after controlling for initial scores.

Measure	F statistic	p-value
GAD-7	$F(2, 171) = 0.11$	0.898
PHQ-9	$F(2, 169) = 0.22$	0.802
PCL-5	$F(2, 160) = 0.07$	0.935

In summary, while clients with different initial ITQ diagnoses may present with different initial severity levels, their trajectories during therapy are equivalent. Final scores are well predicted by a simple linear function of initial scores, with a slope of approximately 0.4–0.5 across all measures. Once initial severity is accounted for, initial ITQ diagnosis provides no additional prognostic information regarding treatment outcomes on these measures.

5 Relationship between ITQ diagnosis and PCL-5 scores

To examine whether PCL-5 scores differ across ITQ diagnostic categories (Complex PTSD, PTSD, and No PTSD), we analyzed data from 239 clients across 671 therapy sessions. Because multiple PCL-5 and ITQ measurements were taken for each client across different therapy sessions, we used a linear mixed effects model to account for the non-independence of observations. Specifically, we included random intercepts for both individual clients and session types to control for systematic differences between clients and variations across different types of therapy sessions.

The analysis revealed clear differences in PCL-5 scores across the three ITQ diagnostic groups (Figure 3). Clients diagnosed with Complex PTSD had an estimated marginal mean PCL-5 score of 52.3 (95% CI: 45.5–59.1), which was substantially above the PCL-5 caseness threshold of 32. Those diagnosed with PTSD (without complex features) had a mean score of 35.6 (95% CI: 28.5–42.7), just above the caseness threshold. Clients with no PTSD diagnosis had a mean score of 26.7 (95% CI: 19.9–33.5), well below the threshold.

All pairwise comparisons were statistically significant ($p < .0001$). Complex PTSD scores were 25.6 points higher than No PTSD scores, 16.7 points higher than PTSD scores, and PTSD scores were 8.9 points higher than No PTSD scores. These findings indicate that PCL-5 scores effectively distinguish between ITQ diagnostic categories, with a clear hierarchical pattern: Complex PTSD > PTSD > No PTSD.

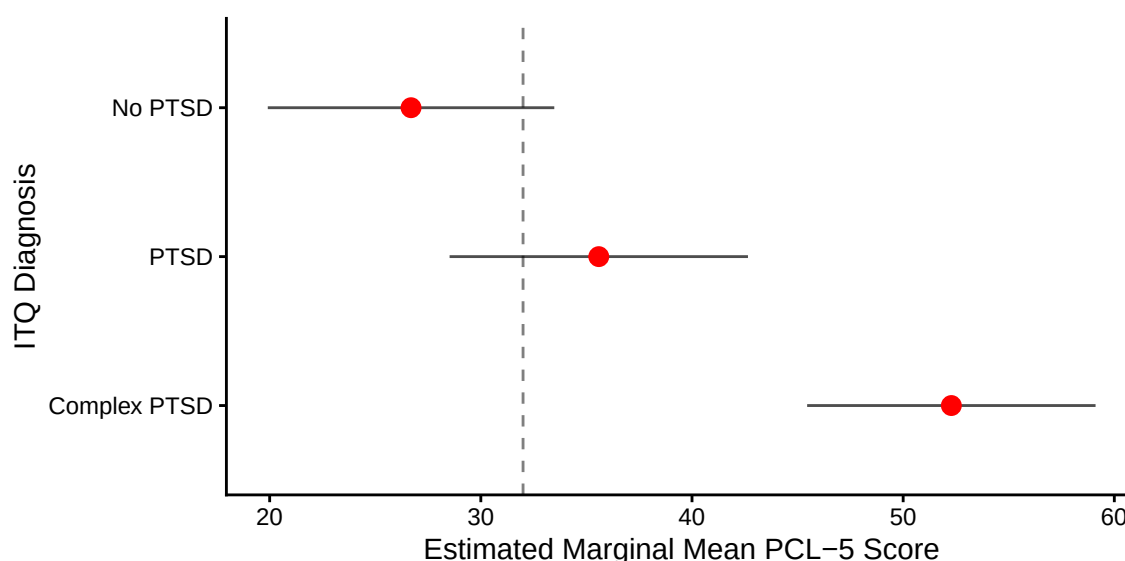


Figure 3: Estimated marginal mean PCL-5 scores by ITQ diagnosis category. Error bars represent 95% confidence intervals. The dashed vertical line indicates the PCL-5 caseness threshold of 32.

6 Executive Summary

This report presents treatment outcome data from 335 clients of PTSD Resolution, a UK charity providing mental health therapy to veterans, reservists, and their families. The analysis reveals consistently positive results across a range of psychological measures, with high engagement rates, significant clinical change, and strong recovery outcomes.

Key findings include:

- **High treatment engagement:** 96.7% of clients (321 of 332) progressed into treatment following screening, receiving one or more therapy sessions with one of 73 therapists. The median number of sessions was 6, with 90% of clients completing between 1 and 12 sessions.
- **Substantial clinical improvement:** Large effect sizes were recorded across all measures (Cohen's $d = 1.0-1.14$), demonstrating meaningful therapeutic change. Reliable improvement rates ranged from 57% to 69% across GAD-7, PHQ-9, and PCL-5 measures.
- **Strong recovery rates:** 40% of clients achieved recovery on combined measures of anxiety (GAD-7) and depression (PHQ-9). On the PCL-5, 53% of initial cases recovered and 69% showed reliable improvement.

The International Trauma Questionnaire (ITQ) provided further diagnostic insight:

- Among clients meeting criteria for PTSD at intake, 90.5% no longer met criteria at discharge.
- Among those initially diagnosed with Complex PTSD, 52.9% transitioned to a "No PTSD" classification, while 36.5% remained in the Complex PTSD category.
- These differences reflect initial severity rather than reduced responsiveness: ANCOVA analyses demonstrated that once baseline severity is controlled, treatment response trajectories were equivalent across diagnostic groups (all $p > .80$).

Furthermore, there is a clear relationship between PCL-5 scores and ITQ classification. Clients with Complex PTSD had an estimated marginal mean PCL-5 score of 52.3, compared to 35.6 for PTSD and 26.7 for No PTSD. All pairwise differences were statistically significant ($p < .0001$), confirming the discriminative validity of the PCL-5 as a predictor of ITQ-defined diagnostic categories.

Logistic regression controlling for initial symptom severity identified living situation and referral source as modest predictors of recovery. Clients living with parents showed the highest estimated probability of recovery (0.97), potentially reflecting social support benefits. Other demographic factors such as age, gender, employment status, or medication use did not significantly impact outcomes once baseline severity was accounted for.

These results provide robust evidence of the effectiveness of PTSD Resolution's therapeutic model, including for clients with Complex PTSD. The use of dimensional and diagnostic tools such as the PCL-5 and ITQ allows precise tracking of clinical change and reinforces the charity's commitment to transparent, evidence-based care.

References

- Cloitre, M., Shevlin, M., Brewin, C. R., Bisson, J. I., Roberts, N. P., Maercker, A., Karatzias, T., & Hyland, P. (2018). The international trauma questionnaire: Development of a self-report measure of ICD-11 PTSD and complex PTSD. *Acta Psychiatrica Scandinavica*, 138(6), 536–546. <https://doi.org/10.1111/acps.12956>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge Academic.
- NHS England. (2019). The improving access to psychological therapies manual: Appendices and helpful resources. *NHS Digital*. <https://www.rcpsych.ac.uk/docs/default-source/improving-care/nccmh/lapt/nccmh-lapt-manual-appendices-helpful-resources-v2.pdf>