
Article:
Robinson, L., Kellett, S. and Delgadillo, J. orcid.org/0000-0001-5349-230X (2020) Dose-response patterns in low and high intensity cognitive behavioral therapy for common mental health problems. Depression and Anxiety, 37 (3). pp. 285-294. ISSN 1091-4269

https://doi.org/10.1002/da.22999

Citation: Robinson, L., Kellett, S., & Delgadillo, J. (*in press*). Dose-response patterns in low and high intensity cognitive behavioural therapy for common mental health problems. *Depression and Anxiety*.

Dose-response patterns in low and high intensity cognitive behavioural therapy for common mental health problems

Louisa Robinson[1], Stephen Kellett[2], & Jaime Delgadillo[2*]

1. Department of Psychology, University of Sheffield, 1 Vicar Lane, Sheffield S1 2LT, UK

2. Clinical Psychology Unit, Department of Psychology, University of Sheffield, 1 Vicar Lane, Sheffield S1 2LT, UK

***Correspondence:** jaime.delgadillo@nhs.net

Abstract

Background: Cognitive behavioural therapy (CBT) is effective for the treatment of common mental health problems, but the number of sessions required to maximise improvement in routine care remains unclear.

Aim: This study aimed to examine the dose-response effect in low (LiCBT) and high (HiCBT) intensity CBT delivered in stepped care services.

Methods: A multi-service dataset included N = 102206 patients across N=16 services. The study included patients with case-level depression and/or anxiety symptoms who accessed LiCBT and/or HiCBT. Patients with post-treatment reliable and clinically significant improvement (RCSI) in standardised outcome measures (PHQ-9, GAD-7) were classified as treatment responders. Survival analyses assessed the number of sessions necessary to detect 50%, 75% and 95% of treatment responders. The 50% and 95% percentiles were used to define the lower and upper boundaries of an *adequate dose* of therapy that could be used to inform the timing of treatment progress reviews. Analyses were then stratified by diagnosis, and cox regression was used to identify predictors of time-to-remission.

Results: Most responders (95%) attained RCSI within 7 sessions of LiCBT and 14 sessions of HiCBT. Patients with social anxiety disorder, post-traumatic stress disorder, and obsessive-compulsive disorder required HiCBT and lengthier treatments (6–16 sessions) to maximise improvement.

Conclusions: Distinctive dose-response patterns are evident for LiCBT and HiCBT, which can be used to support treatment planning and routine outcome monitoring.

Key terms: dose-response; CBT; cognitive behavioural therapy; depression; anxiety

1. INTRODUCTION

The duration and associated costs of psychological treatment vary considerably in routine care. Taking an evidence-based approach, contemporary treatment guidelines for common mental health problems recommend a set number of sessions that are guided by efficacy trials (National Institute for Health and Care Excellence [NICE], 2011). However, such an approach is problematic because we cannot assume that trial participants are representative of patients encountered in routine care, or that all patients have a uniform response to the number of sessions that are set in clinical trials. Furthermore, clinical trial outcomes are usually aggregated at a group level (e.g., intervention cases vs. controls) and assessed at predefined endpoints (e.g., 6-months post-randomization), which precludes the investigation of differential response-times across different patients. For these reasons, it has been argued that practice-based studies with typically variable treatment durations, rather than controlled trials with arbitrary durations, are necessary to better understand the natural variability in response to therapy and to establish pragmatic recommendations for the optimal duration of treatment (Hansen, Lambert, & Forman, 2002; Robinson, Delgadillo, & Kellett, *in press*).

Numerous practice-based studies since the 1980s have found associations between the duration of psychotherapy and symptomatic improvements (Hansen et al., 2002). This *dose-response effect* is characterised by a curvilinear relationship, whereby most of the symptomatic improvement is observed during the earlier stages of treatment (Howard, Kopta, Krause, & Orlinsky, 1986). Researchers have sought to identify the point at which at least 50% of patients who attain clinically significant change are detected, and after which the probability of symptomatic improvement steeply declines (Hansen et al., 2002). Combining the 50% percentile and a higher percentile such as 75% or 95% would enable clinicians to determine *adequate dose* parameters, denoting the minimally acceptable number of sessions and a rational upper limit to decide if a patient is likely to respond to treatment with additional sessions

(Hansen et al., 2002; Howard et al., 1986; Robinson et al., *in press*). This approach recognises that patients' trajectories of change are heterogeneous, since some respond quickly and others require lengthier interventions. However, a cause-effect relationship between treatment dose and clinical outcomes cannot be inferred from uncontrolled studies, since response rates are likely to be influenced by spontaneous remission. Nevertheless, the pragmatic utility of the dose-response model rests in the observation that most patients who attain remission of symptoms (due to therapy, regression to the mean, or spontaneous recovery) show signs of improvement by the upper boundary of the adequate dose.

A systematic review of dose-response studies found that adequate dose recommendations varied according to clinical populations and methodological approach (Robinson et al., *in press*). For example, patients with psychosis required 8–30 sessions of high intensity cognitive behavioural therapy (CBT) (Falkenström, Josefsson, Berggren, & Holmqvist, 2016; Lincoln, Jung, Weisjahn, & Schlier, 2016), while patients with mild-to-moderate anxiety and depression required only 4–6 sessions of low intensity CBT (Delgadillo et al., 2014; Delgadillo, Kellett, et al., 2016). Furthermore, studies using more or less conservative definitions of treatment response yield different conclusions (Anderson & Lambert, 2001; Asay, Lambert, Gregersen, & Goates, 2002). Overall, methodological reviews indicate that adequate dose parameters need to be calibrated in diagnostically homogenous groups, in large and adequately powered samples, using stringent definitions of treatment response (Robinson et al., *in press*).

To date, no studies have examined the dose-response effect for CBT interventions using a sufficiently large sample and appropriate methods to examine patterns in diagnostically homogeneous groups. The present study aimed to address this gap in the literature through the analysis of a large multi-service dataset of low and high intensity CBT delivered in stepped care psychological services. The main objective was to determine an adequate dose of CBT for

different diagnostic groups, and a secondary objective was to identify predictors of time-to-remission.

2. METHODS

2.1. Setting and interventions

This study was conducted using multi-service archival data from 8 National Health Service (NHS) trusts in England, collected between 2014 and 2017. Together, these NHS trusts managed 16 *Improving Access to Psychological Therapies* (IAPT) services, covering socio-economically diverse regions of England; including London, Cambridge, Cheshire & Wirral, Bury, Heywood, Middleton, Rochdale, Oldham, Stockport, Tameside & Glossop, Trafford, Barnsley and East Riding. Verbal consent was obtained and recorded in patients' clinical records to collect weekly symptom measures to enable routine outcome monitoring. The assembly and analysis of a fully anonymised dataset was approved by the London - City & East NHS Research Ethics Committee (06/01/2016, Ref: 15/LO/2200).

IAPT services deliver time-limited evidence-based psychological interventions for depression and anxiety organised in a stepped care model (Clark, 2018; NICE, 2011). Most patients initially access low intensity CBT (LiCBT), which is a brief (usually up to 8 sessions) guided self-help intervention based on principles of CBT. In the present sample, 90% of LiCBT interventions involved individual support, ~8% were delivered in groups, and ~2% delivered by blended care (online CBT plus telephone support). LiCBT is highly structured and follows treatment protocols based on a national training curriculum for psychological wellbeing practitioners (PWPs) (National IAPT Team, 2015) who practice under weekly supervision led by senior practitioners.

Patients who do not respond to LiCBT are "stepped-up" to high intensity psychotherapies. For patients with major depressive disorder, available high intensity

treatments include CBT, interpersonal psychotherapy, short-term dynamic psychotherapy and person-centred counselling. Patients with post-traumatic stress disorder can access CBT or eye-movement desensitization and reprocessing (EMDR). For other common mental disorders, high intensity CBT (HiCBT) is the most commonly recommended first line treatment. HiCBT involves structured, protocol-driven, disorder-specific interventions listed in the Roth & Pilling (2008) competency framework (usually up to 20 sessions). HiCBT was delivered by postgraduate-level psychotherapists whose training was based on a national curriculum (Department of Health, 2011). HiCBT therapists practice under regular clinical supervision (equivalent of 1 hour per week).

Treatment duration in IAPT services can be highly variable, as it is collaboratively agreed with patients and supervisors, it is sometimes curtailed by dropout, and it can therefore differ from the number of sessions recommended in clinical guidelines. In the participating services, 27.6% of patients who started treatment only attended up to 3 sessions. Furthermore, 5% of patients attended more treatment sessions than is specified by clinical guidelines (NICE, 2011), and some (1%) accessed more than twice the number of recommended sessions. In this sample, 59.5% of patients only accessed low intensity interventions and 40.5% accessed high intensity interventions (of whom 17.1% were stepped up after LiCBT, and 23.5% were directly allocated to high intensity). Treatment recovery rates (as defined by Clark 2018) across these services ranged between 40.0% and 52.5%, which were broadly consistent with national performance indices reported by the IAPT programme during the above period (44.8% to 49.3%)*.

2.2. Measures

IAPT services collect outcome measures on a session-by-session basis to monitor treatment progress. The PHQ-9 is a measure of depression symptoms, where each of 9 questions is rated

* National IAPT performance reports are publicly available at: https://digital.nhs.uk/

from 0 to 3, yielding an overall severity score between 0 and 27 (Kroenke, Spitzer, & Williams, 2001). A cut-off of ≥10 has been recommended as providing the best trade-off between sensitivity (88%) and specificity (88%) for a diagnosis of major depression (Kroenke et al., 2001), and a difference of ≥6 points between measurements is indicative of statistically reliable change (Richards & Borglin, 2011). The GAD-7 is a 7-item questionnaire used to identify anxiety disorders; each item is also rated between 0 and 3, with a total severity score between 0 and 21 (Kroenke, Spitzer, Williams, Monahan, & Löwe, 2007). A cut-off score ≥8 is recommended to identify clinically important anxiety symptoms, with adequate sensitivity (77%) and specificity (82%) (Kroenke et al., 2007). A change of ≥5 points has been recommended to assess reliable change (Richards & Borglin, 2011).

Secondary data: De-identified clinical records captured demographics (age, gender, disability, employment status), and clinical information (primary diagnoses, functional impairment, number of treatment sessions received at each step of care). Functional impairment was assessed using the Work and Social Adjustment Scale (WSAS) (Mundt et al., 2002) which rates overall functioning across 5 domains including: work, home management, social life, private leisure activities, and family relationships.

2.3. Case selection process and sample characteristics

In total, 146078 patients accessed treatment across participating services, including those who completed and those who dropped out. The majority of these (N = 116814; 80%) accessed LiCBT and/or HiCBT, with the remainder accessing other psychotherapies. The study sample (N = 102026) only included cases that accessed LiCBT and/or HiCBT and started treatment with clinically significant depression or anxiety symptoms (i.e. baseline scores above the diagnostic cut-offs). The rationale for selecting only low and high intensity CBT cases was twofold: these were the most widely accessed treatments, and other treatments had restricted sample sizes that precluded the detailed survival analyses described below. The study sample

included N = 76962 LiCBT cases and N = 36641 HiCBT cases (N = 12116 had LiCBT + HiCBT; N = 24525 only accessed HiCBT). The characteristics of the sample are summarised in Table 1.

[Table1]

2.4. Statistical analyses

Outcome definition: Sessional data enabled the identification of the first session at which a patient met criteria for reliable and clinically significant improvement (RCSI) on each of the outcome measures. RCSI was recorded (coded = 1) separately for each outcome measure, when scores improved by a magnitude greater or equal to the reliable change index compared to the initial treatment session, and if the score reduced below the diagnostic cut-off (Jacobson & Truax, 1991). The first observed RCSI signal was the primary event of interest in the time-to-event analyses (event = remission of symptoms). All cases were also classified into two groups according to their RCSI outcome status at their final treatment session: *responders* (final session RCSI = 1) and *non-responders* (final session RCSI = 0). In this way, cases that had a short-lived improvement that was not maintained by the end of treatment (referred to as *backsliding*) (Robinson et al., *in press*) could be identified.

Dose-response: Survival analyses identified the number of sessions required for 50%, 75% and 95% of responders to attain remission of symptoms (RCSI). Kaplan-Meier curves were plotted separately for responders and non-responders to identify the stage of treatment at which these groups were reliably differentiated, expecting that some "false positive" cases may show initial signs of improvement that was not maintained to the end of treatment. Separate LiCBT and HiCBT survival analyses were run for each outcome measure. Cases that were stepped up were included separately in each of the subsets of cases according to the step of

treatment (i.e. their data contributed to both sets of analyses). Cases that did not show RCSI at all were included in survival analysis, and they contributed outcomes data up to the point of their last attended session (at which point their data was *censored*). Curves were plotted up to the time point (i.e. session) where there were at least N = 100 (non-censored) cases, to avoid misinterpretation of estimates yielded from small samples. The adequate dose was defined as the interval between 50% and 95% percentiles, where the lower boundary represents the minimum recommended number of sessions, and the upper boundary marks the point after which the probability of response to treatment was negligible (<5% probability). In cases where a diagnosis was recorded in clinical records, the above procedure was run for each diagnostic subgroup estimating post-treatment response rates (RCSI), mean and median survival times (with 95% confidence intervals).

Time-to-remission: Cox regressions identified variables that might influence the relationship between treatment duration and outcomes. Candidate predictors were determined *a priori*, informed by previous outcome-prediction studies in stepped care psychological services (Delgadillo et al., 2014; Delgadillo, Huey, Bennett, & McMillan, 2017; Delgadillo, Moreea, & Lutz, 2016). The regression models for HiCBT cases additionally included a variable that contrasted cases that did or did not access prior LiCBT in their treatment pathway, to assess if accessing consecutive stepped care interventions influenced the dose-response findings. Cox regressions were applied in a dataset where missing values in candidate predictor variables were imputed by aggregating 25 iterations using the Markov Chain Monte Carlo method (Schunk, 2008). Diagnostic categories were not imputed because of the high number of sparsely populated categories; therefore, this analysis was carried out using a subset (N = 73542) of the sample included in the primary survival analysis.

Treatment duration (number of high intensity treatment sessions) and outcomes (RCSI) for HiCBT were compared between cases that did and did not have prior LiCBT, using Mann-Whitney U tests and Chi-square analysis.

3. RESULTS

3.1. Response (RCSI) rates

Overall, between 25.3% (PHQ-9) and 26.8% (GAD-7) of LiCBT cases were classed as responders by the end of treatment. Approximately 4.6% had false-positive RCSI signals before the end of treatment. Response rates were highly variable across the diagnostic groups treated with LiCBT (see Table 2 and supplementary appendix for further details). In particular, the response rates for social anxiety disorder, OCD, somatoform disorder and PTSD indicated that patients with these diagnoses had a low probability of improvement (< 25%) when treated with LiCBT. The response rates for HiCBT were between 38.6% (PHQ-9) and 39.3% (GAD-7), and were consistent across diagnostic groups. HiCBT cases had a higher proportion of non-responders with false positive RCSI signals (9.5%). Patients with somatoform disorder and specific phobias had the highest response rates (~50%), and those with PTSD had the lowest response rates (~33%). No significant differences in HiCBT response rates were found when comparing those that had prior LiCBT and those that only accesed HiCBT, using depression (38.5% vs. 38.6%; $x^2[1] = 0.02$, $p = .89$) or anxiety measures (39.0% vs. 39.5%; $x^2[1] = 0.80$, $p = .37$).

[Table 2]

3.2. Dose-response patterns in LiCBT and HiCBT

Figure 1 presents Kaplan-Meier survival plots, modelling the number of sessions required to identify 50%, 75% and 95% of treatment responders. Findings were consistent across both outcome measures (PHQ-9, GAD-7). For LiCBT, 50% of responders were identified by session 4, 75% by session 5, and 95% by session 7. The cumulative hazard function indicated that cases attaining RCSI by the 7th session were 3 times more likely to be classified as treatment responders. For HiCBT, 50% of responders were identified by session 5, 75% by session 8, and 95% by session 14. The cumulative hazard function indicated that cases attaining RCSI by the 14th session were 3 times more likely to be classified as treatment responders.

[Figure 1]

3.3. Dose-response by diagnostic group

Table 2 summarises the dose-response parameters (50% to 95% boundaries), mean and median survival time estimates in each diagnostic group. For LiCBT, the minimum dose to identify 50% of treatment responders ranged between 3–4 sessions, and 95% of responders were identified by sessions 6–8. The overlapping confidence intervals for mean survival times across groups revealed a highly homogeneous dose-response effect. For HiCBT, the minimum dose to identify 50% of treatment responders ranged between 4–6 sessions, and 95% of responders were identified by sessions 12–16. Mean survival times across groups were also highly homogeneous. The mean survival time for PTSD cases to respond (i.e. 6–16 sessions) was longer than most other conditions, apart from OCD. Specific phobias and somatoform disorders had lower mean survival times compared to other conditions; between 4–13 sessions were required to observe a response in 50% and 95% of cases.

3.4. Predictors of time-to-remission

Table 3 summarises the results of fully adjusted Cox regression analyses. Patients with higher initial impairment (PHQ-9, GAD-7, WSAS), those prescribed antidepressants, those who were unemployed, and those who reported a disability had a decreased probability of attaining RCSI with increasing treatment duration (Hazard Ratio [HR] ~0.65 to 0.98). Older patients were statistically more likely to respond to lengthier treatments, although the effect size was negligible (HR ~1.00). Patients from a minority ethnic group and those with long-term physical health conditions were less likely to attain remission of anxiety symptoms (GAD-7) in lengthier LiCBT; although the effect sizes were small (HR ~0.95). In comparison to depression, patients with social anxiety disorder, OCD and PTSD had a lower probability of response with lengthier interventions (see hazard ratios in Table 3). Furthermore, patients who accessed prior LiCBT tended to access a slightly higher mean number of HiCBT treatment sessions (8.80, SD = 5.37) compared to those who only accessed HiCBT (8.06, SD = 5.62); U(31499) = 116848963.00, $p < .001$. Yet, cases that accessed the full stepped care pathway had a significantly lower probability of response with lengthier HiCBT interventions, compared to those who only accessed HiCBT (HR ~0.87).

[Table 3]

4. Discussion

4.1. Summary of findings

This study investigated the dose-response effect in evidence-based LiCBT and HiCBT interventions for common mental health problems treated in stepped care psychological services. The main findings indicate that the majority (95%) of LiCBT patients who attain reliable and clinically significant improvement (RCSI) do so within 7 sessions, and most HiCBT patients who attain RCSI do so within 14 sessions. These findings are consistent with the wider dose-response literature. Previous LiCBT studies have suggested an adequate dose

of 4–6 sessions (Delgadillo et al., 2014; Delgadillo, Kellett, et al., 2016). Other studies in primary care services which included HiCBT (among other treatments) have found an upper boundary ranging between 11 and 14 sessions (Howard et al., 1986; Wolgast, Lambert, & Puschner, 2003). As expected, predictors of *time-to-remission* were highly consistent with the prognostic factors highlighted by previous studies in similar services (Delgadillo, Kellett, et al., 2016; Delgadillo et al., 2017; Delgadillo, Moreea, et al., 2016). LiCBT adequate dose parameters were homogeneous across diagnostic groups; however, patients with social anxiety disorder, OCD and PTSD had poor response rates (<25%). Response rates for these diagnoses were considerably better when treated with HiCBT; but OCD and PTSD required more sessions (adequate dose of 6–16 sessions).

These adequate dose parameters for LiCBT and HiCBT suggest important differences in response patterns. This is likely to be explained by differences in case-mix, since patients accessing HiCBT tend to have more severe conditions. The evident structural differences between treatments (e.g. LiCBT psychoeducational approach vs. HiCBT psychotherapeutic approach; 35 vs. 50-60 minute sessions) may also explain the differential response rates observed between treatments. These findings highlight the importance of considering dose-response parameters alongside overall response rates. Otherwise, LiCBT may appear to be a more efficient option for the treatment of those diagnoses which clearly require HiCBT. These findings emphasise the importance of adhering to clinical guidelines that recommend HiCBT for social anxiety disorder and PTSD (NICE, 2011). The present results also indicate that HiCBT is a more appropriate treatment choice for OCD.

Patients who accessed the full stepped care pathway (LiCBT+HiCBT) tended to access lengthier interventions, but had similar response rates to those who were directly allocated to HiCBT. An interesting implication is that exposure to prior LiCBT does not apparently confer any advantages (e.g. preparation, socialisation, early gains) for these cases – since their HiCBT

is neither briefer nor more effective. Furthermore, LiCBT+HiCBT cases were statistically less likely to respond to treatment with additional sessions. This latter finding may be explained by the influence of non-responders that persist with treatment (i.e. they have not dropped out early, which is why they eventually access HiCBT), but who nevertheless have a low probability of improvement with the treatments available in this stepped care context.

We note that the response rates described in this study (ranging between 25.3% and 39.3%) are markedly different to recovery rates that are publicly reported for IAPT services in England – which average around 50% (Clark, 2018; National Collaborating Centre for Mental Health, 2018). These differences are due to three methodological features. First, compared to IAPT *recovery rates* which classify cases based on symptomatic reductions below the cut-offs for PHQ-9 and GAD-7 (National Collaborating Centre for Mental Health, 2018), the strict RCSI criteria applied in this study additionally requires the observation of statistically reliable improvement and therefore yields more conservative estimates of treatment response. As shown in a methodological review on this topic (Robinson et al., *in press*), studies that use RCSI criteria (rather than more lenient indices of improvement) tend to yield lengthier dose-response parameters which are less likely to classify cases that take longer to benefit from therapy as non-responders. Second, we report RCSI rates separately for low and high intensity interventions, whereas IAPT recovery rates reflect improvements observed during the whole stepped care treatment pathway starting with an initial (pre-treatment) assessment. Third, the response rates in this study are not confounded by pre-treatment changes that occur between the initial assessment and the first therapy session, since survival analyses take the first treatment session as a baseline measure for the RCSI calculation. Previous studies have shown that pre-treatment improvements (which cannot be attributed to therapy) significantly influence patients' probability of symptom remission (Delgadillo et al., 2014) and must therefore be controlled in dose-response analyses. Finally, the RCSI rates observed for HiCBT are

consistent with the effects observed in meta-analyses of controlled trials, which typically report a number needed to treat (NNT) of 2.6 (~38.5% response rate) (Cuijpers et al., 2013). However, the RCSI rates for LiCBT were considerably lower, possibly owing to the inclusion of cases with conditions like social anxiety and PTSD, which evidently do not respond well to brief guided self-help, and should not be assigned to LiCBT according to clinical guidelines (National Collaborating Centre for Mental Health, 2018; NICE, 2011). Overall, the results presented in this study use a rigorous and conservative methodology, which follows best-practice recommendations for dose-response research (Robinson et al., *in press*).

4.2. Limitations

A number of limitations are relevant to the interpretation of findings. As a naturalistic study, patients were not randomly allocated to treatments and therefore outcomes cannot be compared as if their characteristics were equally balanced across LiCBT and HiCBT. For this reason, the steps were analysed separately, in order to reflect the dose-response patterns observed in these naturally clustering samples of patients with milder and more severe problems. Diagnostic groupings were based on semi-structured assessments which may have produced inaccurate categorisation, so the comparisons across diagnostic groups should be interpreted with caution. PHQ-9 and GAD-7 measures were used to examine change across all diagnostic categories; whereas the availability of disorder-specific measures (i.e. specific measures for PTSD, OCD, etc.) may have provided more accurate indices of improvement.

An important consideration is that, in the absence of a control group, it cannot be assumed that the remission of symptoms was necessarily due to the action of therapy. It is plausible that some patients may have experienced spontaneous remission of symptoms. Nevertheless, this does not negate the relevance of the adequate dose concept, since it still offers a general guide as to when clinicians may expect to observe most cases with remission of symptoms (due to the effects of treatment or natural recovery). Finally, some cases treated

with LiCBT may have been discharged shortly after the first observation of symptom remission. It is possible that some of these cases may have been incorrectly identified as responders, since post-treatment follow-up data were not available to determine the stability of patients' remission over time.

4.3. Implications for clinical practice and policy

Since the 1980's, it has been recognised that some CBT patients attain rapid improvements during the initial therapy sessions (Rush, Kovacs, Beck, Weissenburger, & Hollon, 1981), which suggests that those who respond to therapy can be either *rapid* or *gradual* responders (Robinson et al., *in press*). The first group show signs of improvement within the first 4 sessions, which is recommended as the minimum number of sessions that should be offered in routine care. However, gradual responders require more sessions to benefit and tend to require HiCBT. Identifying the characteristics of these gradual responders (e.g. more complex cases) could inform the development of stratified models of care, in which such cases are immediately assigned to more intensive/longer treatments (Delgadillo et al., 2017; Delgadillo, Moreea, et al., 2016). The case for stratified care is further supported by the present results, since the probability of improvement was comparable between cases that were directly allocated to HiCBT (stratified care) and cases that accessed the full stepped care pathway (which is lengthier and costly). A more consistent and targeted application of stratified care for *gradual responders* could potentially improve the cost-effectiveness of treatment in this context.

Identifying adequate dose parameters offers a useful guide for therapists to support routine outcome monitoring. The 50% percentile (4th session of LiCBT; 5th session of HiCBT) is a key marker to plan a treatment review, assessing the case formulation, expectations, agreement on goals and tasks; particularly for non-improving patients to promote collaboration and adherence. The 75th percentile (5th session of LiCBT; 8th session of HiCBT) is a key time-

point to identify obstacles to improvement, since patients who have not attained remission of symptoms by this point have a low probability (<25%) of benefitting from treatment. This should prompt consultation with a clinical supervisor in order to identify and formulate possible obstacles to improvement and to adjust the treatment plan. If no improvement is observed by the 95th percentile (7th session of LiCBT; 14th session of HiCBT), this marks an appropriate time to consider alternative treatment options. From an ethical and health economic point of view, continuing a treatment from which no benefit is being gained could be seen as a form of neglect, an inadequate use of limited healthcare resources, or at least as an opportunity cost, since the patient could have instead accessed other treatment options much sooner.

Conversely, extending therapy beyond these dose-response intervals (up to 24 sessions) (Robinson et al., *in press*) is warranted for patients who have shown signs of RCSI, and we strongly caution against the arbitrary restriction of treatment sessions for these cases. Prior research has shown that IAPT services that offer a low mean number of treatment sessions tend to attain poorer clinical outcomes (Clark et al., 2018), and therefore offering an adequate dose of therapy is central to effective and ethical practice (National Collaborating Centre for Mental Health, 2018). Although observing remission of symptoms is an important milestone in therapy, it is certainly not the optimal endpoint of treatment, since it is important to ensure that symptoms stabilise in the subclinical range before completing treatment. Our findings clearly show that some non-responders show "false positive" signals of improvement, and therefore observing stable RCSI across more than one measurement point is essential. Furthermore, the initial observation of remission usually marks the beginning of relapse prevention planning and subsequent booster sessions in order to maximize the chances of longer-term sustainability of improvements, as recommended in practice guidelines (National Collaborating Centre for Mental Health, 2018). Overall, these dose-response patterns yield practical recommendations

to guide the timely planning of treatment reviews, and evidence-based decisions to extend the duration of treatment or to consider alternative sources of support.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

REFERENCES

Anderson, E. M., & Lambert, M. J. (2001). A survival analysis of clinically significant change in outpatient psychotherapy. *Journal of Clinical Psychology*, *57*(7), 875–888. https://doi.org/10.1002/jclp.1056

Asay, T. P., Lambert, M. J., Gregersen, A. T., & Goates, M. K. (2002). Using patient-focused research in evaluating treatment outcome in private practice. *Journal of Clinical Psychology*, *58*(10), 1213–1225. https://doi.org/10.1002/jclp.10107

Clark, D. M. (2018). Realizing the Mass Public Benefit of Evidence-Based Psychological Therapies: The IAPT Program. *Annual Review of Clinical Psychology*, *14*(1), 159–183. https://doi.org/10.1146/annurev-clinpsy-050817-084833

Clark, D. M., Canvin, L., Green, J., Layard, R., Pilling, S., & Janecka, M. (2018). Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. *The Lancet*, *391*(10121), 679–686. https://doi.org/10.1016/S0140-6736(17)32133-5

Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry*, *58*(7), 376–385. https://doi.org/10.1177/070674371305800702

Delgadillo, J., Huey, D., Bennett, H., & McMillan, D. (2017). *Case complexity as a guide for psychological treatment selection*. *Journal of Consulting and Clinical Psychology,* 85, 835-853. https://doi.org/10.1037/ccp0000231

Delgadillo, J., Kellett, S., Ali, S., McMillan, D., Barkham, M., Saxon, D., … Lucock, M. (2016). A multi-service practice research network study of large group psychoeducational cognitive behavioural therapy. *Behaviour Research and Therapy*, *87*, 155–161. https://doi.org/10.1016/j.brat.2016.09.010

Delgadillo, J., McMillan, D., Lucock, M., Leach, C., Ali, S., & Gilbody, S. (2014). Early changes, attrition, and dose-response in low intensity psychological interventions. *British Journal of Clinical Psychology*, *53*(1), 114–130. https://doi.org/10.1111/bjc.12031

Delgadillo, J., Moreea, O., & Lutz, W. (2016). Different people respond differently to therapy : A demonstration using patient pro fi ling and risk strati fi cation. *Behaviour Research and Therapy*, *79*, 15–22. https://doi.org/10.1016/j.brat.2016.02.003

Department of Health. (2011). *National curriculum for Cognitive Behavioural Therapy courses, Second edition, updated and revised March 2011*. London: DH.

Falkenström, F., Josefsson, A., Berggren, T., & Holmqvist, R. (2016). How much therapy is enough? Comparing dose-effect and good-enough models in two different settings. *Psychotherapy*, *53*(1), 130–139. https://doi.org/10.1037/pst0000039

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services (Structured abstract). *Clinical Psychology: Science and Practice*, *9*(3), 329–343. https://doi.org/10.1093/clipsy.9.3.329

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The Dose-Effect Relationship in Psychotherapy. *American Psychologist*, *41*(2), 159–164. https://doi.org/10.1037/0003-066X.41.2.159

Jacobson, N. S., & Truax, P. (1991). Clinical Significance: A Statistical Approach to Denning Meaningful Change in Psychotherapy Research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (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

Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Löwe, B. (2007). Anxiety

Disorders in Primary Care: Prevalence, Impairment, Comorbidity, and Detection. *Annals of Internal Medicine*, *146*(5), 317–325. https://doi.org/10.7326/0003-4819-146-5-200703060-00004

Lincoln, T. M., Jung, E., Weisjahn, M., & Schlier, B. (2016). What is the minimal dose of cognitive behavior therapy for psychosis? An approximation using repeated assessments over 45 sessions. *European Psychiatry*, *38*, 31–39. https://doi.org/10.1016/j.eurpsy.2016.05.004

Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *British Journal of Psychiatry*, *180*(05), 461–464. https://doi.org/10.1192/bjp.180.5.461

National Collaborating Centre for Mental Health. (2018). *The Improving Access to Psychological Therapies Manual*. Retrieved from https://www.england.nhs.uk/wp-content/uploads/2018/06/the-iapt-manual.pdf

National IAPT Team. (2015). *National curriculum for the education of Psychological Wellbeing Practitioners, Third edition.* London: NHS England/Department of Health.

National Institute for Health and Care Excellence. (2011). Common mental health problems: identification and pathways to care. London: National Collaborating Centre for Mental Health.

Richards, D. A., & Borglin, G. (2011). Implementation of psychological therapies for anxiety and depression in routine practice: Two year prospective cohort study. *Journal of Affective Disorders*, *133*(1–2), 51–60. https://doi.org/10.1016/J.JAD.2011.03.024

Robinson, L., Delgadillo, J., & Kellett, S. (2019). The dose-response effect in routinely delivered psychological therapies: A systematic review. *Psychotherapy Research*. https://doi.org/10.1080/10503307.2019.1566676

Roth, A. D., & Pilling, S. (2008). Using an Evidence-Based Methodology to Identify the

Competences Required to Deliver Effective Cognitive and Behavioural Therapy for Depression and Anxiety Disorders. *Behavioural and Cognitive Psychotherapy*, *36*(02), 129–147. https://doi.org/10.1017/S1352465808004141

Rush, A. J., Kovacs, M., Beck, A. T., Weissenburger, J., & Hollon, S. D. (1981). Differential effects of cognitive therapy and pharmacotherapy on depressive symptoms. *Journal of Affective Disorders*, *3*(3), 221–229. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6456289

Schunk, D. (2008). A Markov chain Monte Carlo algorithm for multiple imputation in large surveys. *AStA Advances in Statistical Analysis*, *92*(1), 101–114. https://doi.org/10.1007/s10182-008-0053-6

Wolgast, B. M., Lambert, M. J., & Puschner, B. (2003). The Dose-Response Relationship at a College Counseling Center. *Journal of College Student Psychotherapy*, *18*(2), 15–29. https://doi.org/10.1300/J035v18n02

Table 1. Sample characteristics

	All cases N = 102026	**LiCBT cases N = 76962**	**HiCBT cases N = 36641**
Demographics			
Male	35026/101613 (34.5%)	26422/76667(34.5%)	12407/36510 (34.0%)
Female	66587/101613 (65.5%)	50245/76667 (65.5%)	24103/36510 (66.0%)
Age (SD)	39.08 (14.58)	39.58 (14.82)	37.60 (13.65)
Unemployed	26195/95515 (27.4%)	18585/71675 (25.9%)	10983/35020 (31.4%)
Ethnicity			
White British	80788/95739 (84.4%)	60726/ 72365 (83.9%)	29213/34360 (85.0%)
Other	14951/95739 (15.6%)	11639 /72365 (16.1%)	5147/34360 (15.0%)
Clinical Characteristics			
Diagnosis			
Affective disorder	34144/90377 (37.8%)	26721 /67992 (39.3%)	11659/33154 (35.2%)
GAD	13902/90377 (15.4%)	10837/ 67992 (15.9%)	4828/33154 (14.6%)
Mixed anx and dep	26572/90377 (29.4%)	21231/ 67992 (31.2%)	7946/33154 (24.0%)
Panic / agoraphobia	3168/90377 (3.5%)	2283/ 67992 (3.4%)	1291/33154 (3.9%)
Social phobia	1877/90377 (2.1%)	953/ 67992 (1.4%)	1259/33154 (3.8%)
Specific phobia	714/90377 (0.8%)	355/ 67992 (0.5%)	451/33154 (1.4%)
OCD	2240/90377 (2.5%)	814/ 67992 (1.2%)	1840/33154 (5.5%)
PTSD	3039/90377 (3.4%)	957/ 67992 (1.4%)	2576/33154 (7.8%)
Somatoform disorder	722/90377 (0.8%)	349/ 67992 (0.5%)	508/33154 (1.5%)
Other	3884/90377 (4.4%)	3492/ 67992 (5.1%)	796/33154 (2.4%)
Baseline PHQ-9 score (SD)	16.21 (5.76)	15.53 (5.6)	16.22 (5.71)
Baseline GAD-7 score (SD)	14.61 (4.54)	14.14 (4.41)	14.93 (4.32)
Baseline WSAS score (SD)	21.31 (9.47)	20.51 (9.33)	21.86 (9.17)
Mean number of sessions (SD, IQR)	7.17 (5.18, 3 – 9)	3.70 (2.76, 1 – 6)	8.33 (5.51, 4 – 12)

Notes: percentages exclude cases with missing data; GAD = generalized anxiety disorder; OCD = obsessive-compulsive disorder; PTSD = post-traumatic stress disorder; PHQ-9 = depression measure; GAD-7 = anxiety measure; WSAS = work and social adjustment measure; SD = standard deviation; IQR = interquartile range

Table 2. Kaplan-Meier survival estimates, stratified by diagnostic category, across intensities of CBT

Low intensity CBT					
Diagnosis	**N = 63933**	**%Responders (95% CI)**	**50–95% boundaries**	**Mean survival time (95% CI)**	**Median survival time (95% CI)**
Affective Disorder	24435	25.2 (24.7, 25.7)	4 – 7	4.09 (4.04, 4.14)	4.00 (3.94, 4.06)
GAD	10465	32.6 (31.7, 33.5)	3 – 7	3.80 (3.74, 3.86)	3.00 (2.92, 3.08)
Mixed	20162	26.9 (26.3, 27.5)	4 – 8	4.14 (4.08, 4.19)	4.00 (3.94, 4.06)
Panic/ Agoraphobia	2212	33.1 (31.2, 35.1)	3 – 7	3.77 (3.63, 3.91)	3.00 (2.83, 3.17)
Social Anxiety Disorder	910	19.3 (16.9, 22.0)	4 – 8	4.36 (4.07, 4.65)	4.00 (3.65, 4.35)
Specific Phobia	350	31.1 (26.5, 36.2)	3 – 7	3.44 (3.14, 3.74)	3.00 (2.66, 3.34)
OCD	803	13.2 (11.0, 15.7)	3 – 8	4.04 (3.63, 4.44)	3.00 (2.44, 3.56)
PTSD	935	4.8 (3.6, 6.4)	4 – 7	4.13 (3.65, 4.62)	4.00 (3.20, 4.81)
Somatoform disorder	343	22.2 (18.1, 26.8)	4 – 7	3.88 (3.45, 4.31)	4.00 (3.47, 4.54)
Other	3318	30.1 (28.6, 31.7)	3 – 8	3.92 (3.79, 4.06)	3.00 (2.86, 3.14)
High intensity CBT					
Diagnosis	**N = 31643**	**%Responders (95% CI)**	**50–95% boundaries**	**Mean survival time (95% CI)**	**Median survival time (95% CI)**
Affective Disorder	10851	37.8 (36.9, 38.7)	5 – 14	6.30 (6.18, 6.41)	5.00 (4.86, 5.15)
GAD	4706	48.1 (46.7, 49.5)	6 – 14	6.40 (6.24,6.56)	6.00 (5.81, 6.19)
Mixed	7639	37.9 (36.8, 39.0)	5 – 14	6.46 (6.32, 6.61)	5.00 (4.82, 5.18)
Panic/ Agoraphobia	1266	40.8 (38.1, 43.5)	5 – 14	6.35 (6.00, 6.69)	5.00 (4.57, 5.43)
Social Anxiety Disorder	1198	42.8 (40.0, 45.6)	6 – 13	6.44 (6.12, 6.75)	6.00 (5.62, 6.38)
Specific Phobia	434	47.5 (42.8, 52.2)	5 – 12	5.35 (4.88, 5.83)	5.00 (4.29, 5.71)
OCD	1796	42.8 (40.5, 45.1)	6 – 16	6.82 (6.50, 7.14)	6.00 (5.65, 6.36)
PTSD	2493	32.0 (30.2, 33.8)	6 – 16	7.20 (6.89, 7.51)	6.00 (5.61, 6.39)
Somatoform disorder	498	55.0 (50.6, 59.3)	4 – 12	5.37 (4.98, 5.77)	4.00 (3.49, 4.51)
Other	762	45.4 (41.9, 49.0)	5 – 16	6.10 (5.65, 6.55)	5.00 (4.61, 5.39)

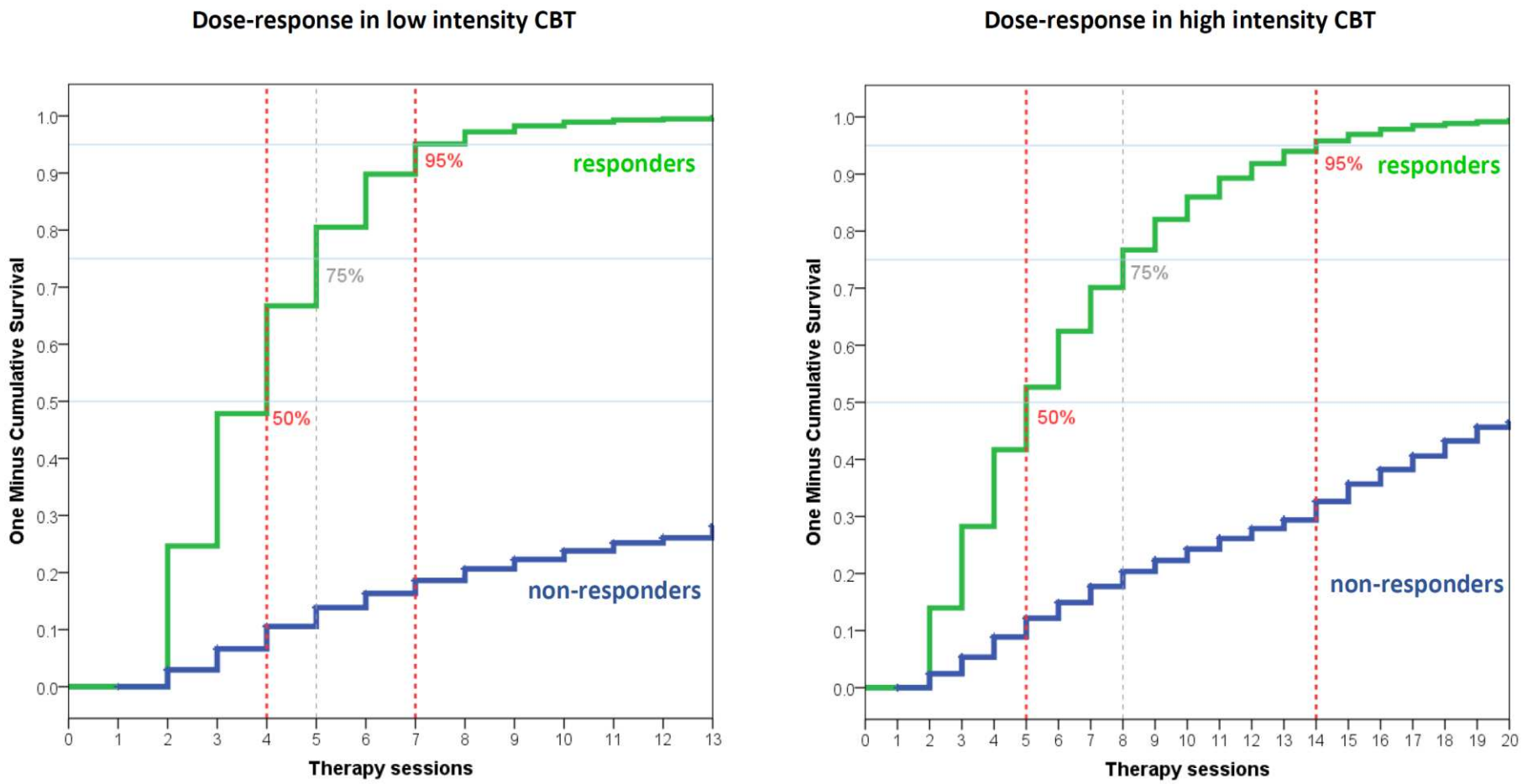
Notes: Responders = % of cases that attained reliable and clinically significant improvement (RCSI) by the end of treatment; 50 – 95% boundaries = the adequate dose of treatment; CI = confidence intervals; GAD= Generalised Anxiety Disorder; Mixed= Mixed anxiety and depression; OCD= Obsessive-Compulsive Disorder; PTSD = Post-Traumatic Stress Disorder; statistics for cases with affective disorders are based on the PHQ-9 depression measure; statistics for cases with anxiety and other disorders are based on the GAD-7 anxiety measure; details of all analyses using both outcome measures in each diagnostic subgroup are available in a supplementary appendix.

Table 3. Cox Regression investigating predictors of time to remission of symptoms

	Low intensity CBT				High intensity CBT			
	PHQ-9 (N = 40790)		GAD-7 (N = 44746)		PHQ-9 (N = 26467)		GAD-7 (N = 28796)	
Variables	**B (SE)**	**Hazard Ratio (95% CI)**	**B (SE)**	**Hazard Ratio (95% CI)**	**B (SE)**	**Hazard Ratio (95% CI)**	**B (SE)**	**Hazard Ratio (95% CI)**
PHQ-9	-0.02(0.002)***	0.98(0.90,0.98)	-0.04(0.002)***	0.96(0.96,0.96)	-0.03(0.002)***	0.97(0.97,0.98)	-0.04(0.002)***	0.96(0.96,0.97)
GAD-7	-0.02(0.002)***	0.98(0.98,0.99)	-0.02(0.002)***	0.99(0.98,0.99)	-0.01(0.002)***	0.99(0.98,0.99)	-0.02(0.002)***	0.98(0.98,0.99)
WSAS	-0.02(0.001)***	0.98(0.98,0.98)	-0.02(0.001)***	0.98(0.98,0.99)	-0.02(0.001)***	0.98(0.98,0.98)	-0.02(0.001)***	0.98(0.98,0.99)
Age	0.004(0.001)***	1.00(1.00,1.01)	0.01(0.001)***	1.01(1.004,1.005)	0.002(0.001)*	1.002(1.001,1.003)	0.003(0.001)***	1.003(1.002,1.004)
MEG	-0.007(0.021)	0.99(0.95,1.03)	-0.04(0.02)*	0.96(0.92,1.00)	-0.03(0.02)	0.97(0.93,1.02)	-0.02(0.02)	0.98(0.94,1.03)
Unemployed	-0.43(0.021)***	0.65(0.62,0.68)	-0.37(0.02)***	0.69(0.66,0.72)	-0.34(0.02)***	0.71(0.68,0.74)	-0.33(0.02)***	0.72(0.69,0.75)
Medication	-0.07(0.015)***	0.93(0.91,0.96)	-0.004(0.01)	0.99(0.97,1.02)	-0.05(0.02)**	0.95(0.92,0.98)	-0.04(0.02)*	0.96(0.93,0.99)
LTC	-0.03(0.018)	0.97(0.94,1.01)	-0.05(0.02)**	0.95(0.92,0.98)	-0.04(0.02)	0.96(0.93,1.00)	-0.04(0.02)	0.97(0.93,1.00)
Disability	-0.23(0.026)***	0.79(0.75,0.83)	-0.17(0.03)***	0.85(0.81,0.89)	-0.15(0.03)***	0.86(0.81,0.91)	-0.10(0.03)***	0.91(0.86,0.96)
Diagnosis †								
(GAD)	0.08(0.023)***	1.09(1.04,1.14)	-0.04(0.02)	0.97(0.93,1.01)	0.06(0.03)*	1.07(1.01,1.12)	-0.06(0.03)*	0.95(0.90,0.99)
(mixed)	-0.02(0.018)	0.98(0.94,1.01)	-0.06(0.02)***	0.94(0.91,0.97)	0.02(0.02)	1.02(0.98,1.07)	0.01(0.02)	1.01(0.96,1.05)
(panic)	0.06(0.043)	1.07(0.98,1.16)	-0.02(0.04)	0.98(0.91,1.05)	-0.04(0.05)	0.96(0.87,1.05)	-0.03(0.04)	0.97(0.89,1.05)
(SocAnx)	-0.20(0.077)*	0.82(0.71,0.96)	-0.31(0.07)***	0.74(0.64,0.85)	0.01(0.05)	1.01(0.92,1.10)	-0.11(0.04)*	0.90(0.83,0.98)
(Phob)	0.28(0.116)*	1.33(1.06,1.67)	-0.07(0.09)	0.94(0.78,1.12)	0.13(0.08)	1.14(0.98,1.32)	-0.03(0.06)	0.98(0.86,1.10)
(OCD)	-0.15(0.109)	0.86(0.70,1.07)	-0.43(0.09)***	0.65(0.55,0.77)	-0.05(0.04)	0.95(0.88,1.03)	-0.27(0.04)***	0.76(0.71,0.82)
(PTSD)	-0.62(0.139)***	0.54(0.41,0.71)	-0.68(0.14)***	0.51(0.39,0.66)	-0.17(0.04)***	0.84(0.79,0.90)	-0.26(0.04)***	0.77(0.72,0.83)
(soma)	0.15(0.125)	1.16(0.91,1.48)	-0.03(0.10)	0.98(0.80,1.19)	0.20(0.07)**	1.22(1.06,1.40)	-0.002(0.06)	0.99(0.88,1.12)
(other)	0.01(0.036)	1.01(0.94,1.09)	-0.02(0.03)	0.98(0.92,1.05)	0.07(0.06)	1.07(0.96,1.19)	-0.01(0.05)	0.99(0.90,1.10)
Prior LiCBT	-	-	-	-	-0.15(0.02)***	0.86(0.83,0.90)	-0.13(0.02)***	0.88(0.85,0.91)

B = regression coefficient; SE = standard error; CI = confidence intervals, PHQ-9= depression measure at initial therapy session; GAD-7= anxiety measure at initial therapy session; WSAS = measure of functional impairment at initial therapy session; MEG = minority ethnic group (reference category = white British); Medication = prescribed antidepressants; LTC = long-term health condition; † = reference category for diagnosis = affective disorders; GAD= Generalised Anxiety Disorder; mixed= Mixed anxiety and depression; panic = panic disorder / agoraphobia; SocAnx = social anxiety disorder; Phob = specific phobia; OCD= Obsessive Compulsive Disorder; PTSD = Post-Traumatic Stress Disorder; soma = somatoform disorder; Prior LiCBT = cases that had low intensity CBT followed by high intensity CBT (versus cases that only had high intensity CBT); *** p <.001; ** p < .05; * p < .05

Figure 1. Kaplan-Meier survival curves comparing treatment responders to non-responders



Notes: Vertical line-markers indicate the number of sessions at which 50%, 75% and 95% of treatment responders are identified in low and high intensity forms of cognitive behavioural therapy.