**Article:**
Robinson, L., Delgadillo, J. and Kellett, S. (2019) The dose-response effect in routinely delivered psychological therapies: A systematic review. Psychotherapy Research, 30 (1). pp. 76-96. ISSN: 1050-3307

https://doi.org/10.1080/10503307.2019.1566676

This is an Accepted Manuscript of an article published by Taylor & Francis in Psychotherapy Research on 20/01/2019, available online: http://www.tandfonline.com/10.1080/10503307.2019.1566676

# The dose-response effect in routinely delivered psychological therapies:

# A systematic review

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

1. Department of Psychology, University of Sheffield, UK

2. Clinical Psychology Unit, Department of Psychology, University of Sheffield, UK

**Correspondence:**

jaime.delgadillo@nhs.net

**Abstract**

The dose-response effect refers to the relationship between the dose (e.g., length, frequency) of treatment and the subsequent probability of improvement. This systematic review aimed to synthesise the literature on the dose-response effect in routine psychological therapies delivered to adult patients with mental health problems. Twenty-six studies were eligible for inclusion. Different methodological approaches have been used to examine the dose-response effect; including survival analysis, multilevel modelling and descriptive cluster analyses. Replicated and consistent support was found for a curvilinear (log-linear or cubic) relationship between treatment length and outcomes, with few exceptions such as eating disorders and severe psychiatric populations. Optimal doses of psychotherapy in routine settings range between 4 – 26 sessions (4 – 6 for low intensity guided self-help) and vary according to setting, clinical population and outcome measures. Weekly therapy appears to accelerate the rate of improvement compared to less frequent schedules. Most of the reviewed evidence is from university counselling centres and outpatient psychotherapy clinics for common mental health problems. There is scarce and inconclusive evidence in clinical samples with chronic and severe mental disorders.

**Key terms**: dose-response, psychotherapy, outcomes research, systematic review

## Introduction

The expression "dose-response" is derived from the pharmacological literature and refers to the relationship between the dose (e.g., quantity and/or concentration) of a treatment and the subsequent probability of improvement. In a psychotherapy context, "dose" commonly refers to the number of therapy sessions delivered, and "response" refers to evidence of clinically and/or statistically defined improvement. In a seminal study, Howard, Kopta, Krause and Orlinsky (1986) first examined the dose-response effect in psychotherapy using data from 2,431 cases across 15 studies. Three important observations were derived from this work. First, the relationship between the number of sessions and treatment outcome conforms to a log-linear function (a negatively accelerating or curvilinear trend). Second, most of the therapeutic change occurs during the earlier sessions of treatment (usually by session 8). Third, this curvilinear relationship predicts diminishing improvements over time. Based on these observations, an "optimal dose of therapy" could be defined as an interval, between the point at which at least 50% of treatment responders are detected, and the point after which response rates plateau (no further improvements observed with additional sessions). The rationale for the 50% response rule is based on the notion that the majority of treatment responders are identified by this point, and the remaining cases have an increasingly low (<50%) probability of improvement thereafter.

Since Howard and colleagues' pioneering work, several studies have examined the dose-response effect in randomised controlled trials (e.g., Barkham et al., 1996) and naturalistic settings (e.g., Hansen, Lambert, & Forman, 2002). Methods used in these studies include survival analysis (e.g. Anderson & Lambert, 2001; Hansen & Lambert, 2003), longitudinal multilevel modelling (e.g., Baldwin, Berkeljon, Atkins, Olsen, & Nielsen, 2009; Reese, Toland, & Hopkins, 2011), logistic regression (e.g., Baldwin et al., 2009; Erekson, Lambert, & Eggett, 2015), cumulative dose-response curves (e.g., Callahan and Hynan, 2005; Delgadillo

et al., 2016) and other descriptive methods. Furthermore, outcome definitions vary across studies, including estimations of clinically significant and/or statistically reliable improvement as proposed by Jacobson and Truax (1991). The dose-response literature is characterised by studies with diverse methodological approaches, patient diagnoses, psychotherapies, outcome measures used, and criteria applied to define improvement. Consequently, inconsistent "optimal dose" recommendations have been proposed in different studies. For example, some studies in university counselling centres (UCCs) suggest an optimal dose around 8 – 11 sessions (Anderson & Lambert, 2001; Baldwin et al., 2009), whilst other studies in UCCs recommend 4 – 10 sessions (Draper et al., 2002). Similar discrepancies are found in studies with primary care clinical samples, where some recommend between 4 – 6 sessions (Delgadillo et al., 2014, 2016) and others recommend 8 – 13 sessions (Falkenstrom et al., 2016).

Several studies have suggested that an alternative perspective, the *good-enough level* (GEL) model, better represents the nature of patient response to psychotherapy (Baldwin et al., 2009; Barkham et al., 1996, 2006; Falkenstrom et al., 2016; Owen et al., 2016; Reese et al., 2011; Stiles et al., 2008, 2015). According to the GEL model, patients who attend different lengths of treatment change at different rates; some respond quickly within a few sessions while others respond gradually with longer treatments. In contrast to the dose-response model, the GEL model proposes that the effect of additional sessions is not (on average) equal across people, and therefore the duration of therapy is unrelated to the probability of improvement or negatively related in samples with a large proportion of rapid responders (Barkham et al., 2006). Thus, there are key theoretical distinctions between these alternative models. The dose-response model assumes that exposure to longer treatments *causes* symptomatic improvements in a similar way across patients, but the strength of each additional session weakens over time. While the GEL model assumes that differential treatment durations *reflect* the extent to which subgroups of patients are responsive to therapy, where longer treatments are a proxy marker

for more difficult-to-treat subgroups. Within the GEL literature, there is mixed evidence regarding the extent to which dose-outcome relationships within subgroups of cases with similar durations of treatment may be linear or curvilinear (Barkham et al., 2006; Baldwin et al., 2009; Reese et al., 2011; Stiles et al., 2008, 2015).

Given this highly heterogeneous literature, the clinical and organisational implications are unclear. The present systematic review therefore had two objectives; the first methodological and the second practical. The first objective was to summarise the methodologies used to explore the relationship between treatment duration and outcomes, in order to understand how the findings may be related to the choice of methods. The second objective was to assess if the current evidence could glean a clear indication of an optimal dose of psychotherapy in routine care, and to define any differences in the optimal length of treatment for different clinical settings or psychotherapies.

[Figure 1]

## Methods

### Protocol and registration

A systematic review protocol was published prospectively in the PROSPERO database. [http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018088886].

### Search strategy and study selection

The following databases were searched (1946 to present): PsychInfo, Scopus, and MEDLINE (including Epub ahead of print, in-process and other non-indexed citations). Key search terms were combined using Boolean operators (AND, OR), and included different synonyms of the terms: dose, length, duration, psychotherapy, outcome, and response. Following the database searches, reference lists of studies eligible for inclusion were hand searched and a forward citations were searched using Google Scholar. Any full texts that could not be sourced were requested by contacting authors. In order to address the objectives of the review, the search

focused on identifying studies which made explicit reference to, and attempted to study, the dose-response effect.

**Inclusion and exclusion criteria**

Studies eligible for inclusion included (1) adults aged 16 and over; (2) patients with psychological disorders and symptoms, receiving routinely delivered therapies; (3) using validated clinical outcome measures. We excluded (1) studies which randomised or allocated participants to treatment conditions; and (2) studies which explored extensions of the dose-response relationship such as the phase model, without reference to dose-response.

The rationale for only reviewing naturalistic studies was two-fold. First, routine care studies often allow variable treatment durations, unlike controlled trials which commonly stipulate minimum and maximum treatment durations (thus limiting the possibility of studying the dose-response effect in variable-length treatments). Second, we chose to study naturalistic samples in order to derive findings that were generalizable and most directly informative for routine psychological care. Figure 2 details the selection process which resulted in the studies eligible for inclusion.

[Figure 2]

**Data Extraction**

Data from eligible studies were collected using a standardised data extraction form. Study characteristics including patient demographics, symptoms and diagnoses, interventions, study settings, primary outcome measures and methodological approaches were the key details extracted, along with findings on the optimal dose and nature of the dose-response effect.

**Quality Assessment**

All studies eligible for inclusion were assessed using the 12-item CASP checklist for cohort studies (Critical Appraisal Skills Programme, 2018). Each question requires “yes”, “can’t tell” or “no” answers which were then subsequently assigned scores of 1, 0 and 0, respectively. A

total score was calculated for each study, out of a maximum of 12 points, where scores <6 are indicative of below average quality. The full set of studies were quality-assessed by the first author, with a random selection of 50% of the studies independently rated by a second assessor. The interrater reliability for this subset of studies was Kappa = 0.865, indicating good agreement.

**Data analysis**

A narrative synthesis of the findings was planned *a priori*, in line with the study objectives, grouping findings by the methods used to study the dose-response effect, followed by a breakdown of the findings by patient groups, therapies and any other relevant groupings identified.

[Table 1]

## Results

**Overview of included studies**

In total, 26 studies met the criteria for inclusion (see Table 1 for a summary), revealing high heterogeneity in terms of setting, patient group and outcome measures used. More than half of these studies (n = 14) were conducted in University training centres or UCCs, while the rest were conducted in primary care, specialist psychotherapy services or in private practice. The majority of studies included patient samples presenting with multiple diagnoses (mixed), receiving a range of therapies. Six studies did not specify the diagnoses or presenting problems that were being treated and ten studies gave little (or no) information regarding what psychotherapies were delivered.

The majority of studies (n = 21) used a single outcome measure. The Outcome Questionnaire (OQ-45; Lambert et al., 1996) was used in 11 studies; four used the Behavioural Health Measure (BHM-20; Kopta & Lowry, 2002), three used the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM; Barkham, Gilbert, Connell, Marshall, &

Twigg, 2005; Evans et al., 2002), one used the Eating Disorders Examination Questionnaire (EDE-Q; Fairburn & Beglin, 1994), one used the Revised Symptom Checklist (SCL-90-R; Derogatis, 1983), and one used the German Questionnaire for the Evaluation of Psychotherapeutic Change Processes (FEP; Lutz et al., 2009). The remaining five studies applied multiple outcome measures (Beail, Kellett, Newman, & Warden, 2007; Delgadillo et al., 2014, 2016; Lincoln, Jung, Weisjahn, & Schlier, 2016; Snell, Mallinckrodt, Hill, & Lambert, 2001). In terms of how change was defined, most studies used a conventional definition of reliable change (RC) or reliable and clinically significant improvement (RCSI), based on the criteria proposed by Jacobson and Truax (1991). However, many different combinations and operationalisations of these criteria were used (see Table 2 for a summary). [Table 2]

**Quality assessment**

All included studies scored 5 and above out of 12 on the CASP checklist. Quality assessment scores for each of the included studies are presented in Table 1 and detailed in a supplementary appendix. Study quality was seen to vary across studies and no apparent pattern of dose-response recommendations was evident according to quality ratings. Only five studies were found to adequately adjust for all potential confounding factors.

**Methodological approaches to studying the dose-response effect**

***Kaplan-Meier procedure.*** Six studies applied survival analysis, following the Kaplan-Meier procedure to predict the number of sessions required to attain RC or RCSI (Anderson & Lambert, 2001; Asay, Lambert, Gregersen, & Goates, 2002; Carr, Saules, Koch, & Waltz, 2017; Harnett, O'Donovan, & Lambert, 2010; Snell et al., 2001; Wolgast, Lambert, & Puschner, 2003). Cox regression was used in conjunction with survival analysis in two studies, adjusting for pre-treatment symptom severity (Asay et al., 2002; Snell et al., 2001).

Results indicated that for 50% of patients to reach RC, between 7 – 42 sessions were required. For 50% of patients to reach RCSI, 11 – 54 sessions were estimated. This large discrepancy can be mostly attributed to the study by Asay et al. (2002). This study predicted 42 and 54 sessions for 50% probability of RC or RCSI, respectively. Specific characteristics of this study most likely cause results to differ: only a single therapist's cohort of patients was used in the study sample, 66% of patients presented with comorbid axis II personality disorders, and no session limits were imposed in this clinic. Lengthier estimates were also derived by Snell et al. (2001); they measured outcomes at an approximate 10-month follow-up and found that the probability of 50% of patients achieving and maintaining RC and RCSI *post-termination* of therapy, was observed at 14 and 16 sessions, respectively.

The five studies from training clinics are largely homogenous in their sample characteristics. Collectively, they suggest that between 7 – 14 sessions were required for 50% to reach RC and between 11 – 19 sessions were required for 50% to reach RCSI. Three of these studies measured outcomes on a session-by-session basis and utilised the full dataset by extracting the first point at which a patient reaches RC or RCSI, which then needed to be maintained until termination of treatment without backsliding (Anderson & Lambert, 2001; Harnett et al., 2010; Wolgast et al., 2003). Findings from these three studies indicate that 8 (n = 1 study) to 10 (n = 2) sessions were required for 50% RC, and 11 (n = 1) to 14 (n = 2) sessions for 50% RCSI.

***Recovery tables.*** Only one study used survival analysis in the form of recovery tables to estimate the treatment length required for the first observation of RCSI to be maintained until follow-up (Kadera et al., 1996). Recovery tables indicated that 44% of cases would reach RCSI in 13 sessions, and 50% by session 16. Across studies using survival analysis, estimates for optimal dose based on patients' first observation of RC range between 8-13 sessions, and

between 11-16 sessions for RCSI. Lengthier doses are observed in patients with more complex presentations (e.g., psychosis) or to maintain improvement until a 10-month follow-up point.

***Logistic regression.*** Logistic or multilevel logistic regression was used in three studies. Baldwin et al. (2009) used logistic regression to estimate the number of sessions required to reach RCSI. The probability of RCSI increased until 8 sessions, with no further improvements beyond this point. Lincoln et al. (2016) mirrored this method, but to specifically examine the dose-response effect in CBT for psychosis. A minimum of 15 sessions was required for the first significant reduction in symptoms, however, based on subsequent diminishing gains, 20 – 30 sessions was suggested as the optimal dose. Erekson et al. (2015) used multilevel logistic regression to examine if session frequency predicted the likelihood of RCSI. Overall, the results indicated that the probability of achieving RCSI or RC were equal regardless of whether treatment was weekly or fortnightly, with more sessions required for equivalent outcomes in fortnightly therapy.

***Probit analysis.*** Probit analysis was used in one study (Kopta et al., 1994) to examine the dose-response effect in a poly-symptomatic sample, grouping symptoms into three latent classes: acute, chronic and characterological. Findings indicated that the median number of sessions for acute symptoms to reach RCSI was 5; 14 sessions was estimated for chronic symptoms; and for characterological symptoms, less than 50% of patients reached RCSI by the end of 52 sessions. Across the most commonly measured symptoms, 50% of patients were expected to reach RCSI by 11 sessions.

***Analysis of variance.*** Analysis of variance (ANOVA) methods were used in three studies. Beail et al. (2007) studied a sample of individuals with intellectual disabilities, being treated for mental health problems with open-ended psychodynamic psychotherapy. ANOVA was used to test for significant differences in pre-post treatment effect sizes between groups that received different lengths of therapy. No significant differences between 8, 16 and 24

weeks of therapy were observed, suggesting that minimal (or no) gains were observed beyond 8 sessions of psychodynamic therapy for patients with intellectual disabilities. Kopta et al. (2014) also used ANOVA to examine differences in pre-post change scores between groups of patients across 23 UCCs, categorised by total lengths of therapy. Post-hoc analyses revealed that longer treatments were associated with greater change scores until 7 – 10 sessions, but smaller and non-significant gains were observed beyond this point. Bell, Waller, Shafran and Delgadillo (2016) found no evidence of a typical curvilinear or linear dose-response association between treatment duration and EDE-Q scores in a sample of outpatients with eating disorders, but patients who showed early improvements (by session 8) were significantly more likely to attain post-treatment RCSI.

***Cumulative dose-response curves.*** Cumulative response-rate curves of patients reaching clinical significance criteria were presented in three studies (Callahan and Hynan, 2005; Delgadillo et al., 2016; Kadera et al., 1996). Comparisons between each of these studies are limited primarily by differences in the restrictiveness of the sampling of patient groups (i.e. full sample, dysfunctional only or those who achieve RCSI), by the outcome criteria of interest and also by the total number of sessions examined in the study (e.g., a maximum of 9 sessions was plotted in the study by Delgadillo et al., 2016, compared to other studies examining time-unlimited therapies). A common trend was for steep initial recovery curves with diminishing gains in the latter stages of therapy; the point of inflexion occurring between 4 – 24 sessions.

***RCSI rates stratified by total treatment sessions.*** Percentages of patients achieving RC and RCSI stratified by the total number of sessions attended were reported in five studies (Draper et al., 2002; Delgadillo et al., 2016; Delgadillo et al., 2014; Kopta et al, 2014; Snell et al., 2001). In contrast to cumulative percentages, stratified percentages represent the proportion of patients terminating in that range of sessions only, rather than a percentage of the full sample. Within these studies, a minimum of four (Draper et al., 2002; Delgadillo et al., 2014) and a

maximum of 8 – 12 sessions (Snell et al., 2001) was suggested as the optimal dose. However, whilst some studies found a clear indication of an optimal dose (Delgadillo et al., 2014, 2016; Kopta et al., 2014) others demonstrated that using this method yielded unclear findings, marked by fluctuations in patient recovery rates over the duration of therapy (Draper et al., 2002; Snell et al., 2001). Such seemingly random fluctuations of response rates in subgroups of cases with the same number of treatment sessions may well be explained by small sample sizes (e.g. under 100 cases) in each subgroup.

***Correlations.*** Correlational analyses were used in three studies to examine the relationship between the percentage of patients recovered and total duration of therapy. Specifically pertaining to RCSI rates stratified by total sessions within these studies, Stiles, Barkham, Connell, & Mellor-Clark (2008) and Stiles et al. (2015), found a negative overall relationship between RCSI rates and treatment duration. The earlier study found this exclusively in UK primary care settings and the latter replicated this across a range of settings including primary and secondary care settings and UCCs, however this relationship was not observed in voluntary and work place sectors. Overall, the negative relationship between treatment dose and outcome was taken as evidence contrary to the dose-response model, and in favour of the "good enough level" (GEL) model (Barkham et al., 1996).

In direct contrast, Stulz et al. (2013) found a positive relationship between duration of therapy and RCSI across 20 UCCs, 4 primary care facilities and 2 private mental health centres in the UK. With all studies using outcome measures designed to consider well-being, symptoms and functioning, there appears to be some inconsistency in the overall relationship between total dosage and RCSI rates. These inconsistencies may result from the limitations of these descriptive methods, which fail to adjust for potential covariates including diagnoses, therapy types, and initial symptom severity. Where studies were marked by similar patient

characteristics and therapies (e.g. Delgadillo et al., 2014, 2016) much more consistent findings were observed.

***Multilevel modelling and latent growth curve modelling.*** A total of eight studies used multilevel modelling techniques (MLM) and one study used latent growth curve modelling (LGCM). One study was specifically focused on CBT for psychotic disorders (Lincoln et al., 2016). Of the 8 remaining studies, seven used data primarily from UCCs, and the eighth used a sample from both a primary care service and a psychiatric outpatient clinic. All eight studies included poly-symptomatic samples and a mixture of therapeutic approaches. Results indicated that 15 sessions were necessary for 50% of cases to respond to treatment, with diminishing improvements beyond 25 sessions.

***Trajectory and shape of change.*** Statistical modelling techniques were used in eight studies to examine if the characteristic log-linear shape of change is found consistently when studying the dose-response effect. Of these, seven studies used longitudinal MLM and compared the model fit for different functions of time (number of sessions). Models with a log-linear function of sessions were found to offer a significantly better fit than a linear function of sessions in three studies (Owen et al., 2016; Sembill et al., 2017; Stulz et al., 2013). Other studies supported cubic functions of sessions (Baldwin et al., 2009; Erekson et al., 2015; Falkenstrom et al., 2016; Reese et al., 2011). In addition, neither Sembill et al. (2017) nor Stulz et al. (2013) compared log-linear functions with cubic, therefore superiority of log-linear over cubic functions cannot be assumed in these studies, whereas Baldwin et al. (2009) and Falkenstrom et al. (2016) did make this comparison and found support for the latter in UCC and primary care samples. In addition, Owen et al. (2015) used multilevel growth mixture modelling to examine different clusters of patients who appeared to have distinct trajectories of change over the course of therapy. Three latent classes were extracted from the session-by-session data collected from 10,854 patients primarily taken from UCC databases. Whilst three

classes were found, class 1 accounted for 75.3% of patients. These patients were typically moderately functioning at the start of treatment and showed initial good response to therapy until around session 5. This was followed by a plateau until around session 11, and then a subsequent phase of improvement. The described trajectory is similar in nature to a cubic function (rather than a log-linear function). However, these models use intensive time-series (session-by-session) data and examine the shape of change *within-subjects*. This is a different analytical approach to typical dose-response curves which compare rates of change *between-subjects* that are grouped according to their total treatment duration.

***Dose-effect vs GEL models.*** MLM was used in five studies to compare the dose-effect versus the GEL model, and one used LGCM (Stulz et al., 2013). Of these, four studies (Baldwin et al., 2009; Falkenstrom et al., 2016; Owen et al., 2016; Reese et al., 2011) compared fit indices of an aggregate (dose-response) model versus a stratified (GEL) model to test this hypothesis. In the aggregate model, model improvement is assessed by adding only the fixed effect of sessions (or a function of sessions), measured on a session-by-session basis. In this model, all patients are assumed to improve at an equivalent rate. In contrast, the individual's total dose is entered as an additional factor in the stratified model, allowing the rate of change to vary by treatment duration. All six studies unanimously found that patients attending different overall lengths of therapy change at different rates, supporting the GEL model. There is a clear trend among these studies for patients that receive lower doses to be characterised by faster rates of improvement.

***Examination of session frequency.*** MLM with frequency of sessions as a fixed effect was used in two studies, in order to examine the impact of session frequency on speed of recovery (Erekson et al., 2015; Reese et al., 2011). Both studies found the fixed effect of session frequency to significantly improve model fit. In addition, Erekson et al. (2015) found that patients attending weekly therapy were significantly more likely to achieve RC and RCSI

sooner than the fortnightly group. Given this evidence, frequency of sessions could be suggested to yield important implications for the dose-response effect, or more importantly for research purposes, may confound results where this is not accounted for.

**Is there an optimal dose in routine treatment settings?**

Optimal doses for individual studies are presented in Table 1. Across all studies, estimates of the optimal dose ranged between 4 – 54 sessions. Excluding studies with samples characterised by particularly chronic or psychotic symptoms (Asay et al., 2002; Lincoln et al., 2016), between 4 – 24 sessions were then estimated.

***University counselling centres (UCCs).*** Of the 14 studies conducted in UCCs, 12 suggested an optimal dose. This was defined in 2 ways; firstly on the basis of 50% of the sample reaching RC or RCSI (Anderson & Lambert, 2001; Carr et al., 2017; Erekson et al., 2015; Harnett et al., 2010; Snell et al, 2001; Wolgast et al., 2003), and secondly on the basis of the number of sessions at which diminishing gains were observed (Baldwin et al., 2009; Callahan & Hynan, 2005; Draper et al., 2002; Kadera, Lambert, & Andrews, 1996; Kopta et al., 2014; Lincoln et al., 2016; Snell et al., 2001). For the former approach, the number of sessions required to reach 50% RC ranged between 7 – 14 sessions and between 11 – 19 sessions to attain 50% RCSI. For the latter approach, the optimal dose ranged between 4 – 30 sessions. If this is estimated excluding cases with psychosis (Lincoln et al., 2016), then the upper boundary is 24 sessions. However, even with the same data set, Snell et al. (2001) suggest different outcomes according to these 2 approaches: 8 – 12 sessions were suggested on the basis of diminishing gains for RC and RCSI, however survival analyses on the same data suggested that 16 sessions would be required for 50% of the sample to achieve RCSI. Whilst these two findings are not mutually exclusive, it highlights some of the reasons for the disparity in findings across studies. Overall, the optimal dose across studies in UCCs appears to range

between 4 – 24 sessions for those with common mental health problems, although up to 30 sessions may be indicated for cases with more severe mental health problems (e.g., psychosis).

***Primary Care services.*** Five studies examined primary care samples. Two studies focused specifically on "low intensity" psychological therapies (guided self-help based on principles of cognitive behavioural therapy) for depression and anxiety (Delgadillo et al., 2014, 2016). Both studies recommended an optimal dose between 4 – 6 sessions; based on the observation that at least 4 sessions are necessary for 50% of cases to attain RCSI and diminishing RCSI rates were observed in treatments longer than 6 sessions. Unlike many studies included in this review, highly structured, standardised and protocol-driven interventions were used in both of these studies. In contrast, Stiles et al. (2008, 2015) and Falkenstrom et al. (2016) studied a poly-symptomatic sample receiving a mixture of "high intensity" psychotherapies (e.g., person-centred, psychodynamic, cognitive behavioural-therapy), and looked for trends across a greater total number of sessions. Stiles et al. (2008, 2015) found a trend for diminishing RCSI rates with lengthier therapies, but did not recommend an optimal dose. Using session-by-session change scores, Falkenstrom et al. (2016) observed greater RCSI in lengthier therapies up to 13 sessions, after which further gains diminished. Overall, primary care populations with mild-to-moderate depression and anxiety symptoms require between 4 – 6 sessions of low intensity guided self-help. Poly-symptomatic primary care populations with more severe symptoms may require up to 13 sessions to attain RCSI.

***Psychiatric outpatients, eating disorder and patients with intellectual disabilities.*** Falkenstrom et al., (2016) found similar outcomes across psychiatric outpatients regardless of their overall treatment length, indicating no dose-response association in this population. Bell et al. (2016) found no association between treatment duration and pre-post therapy change scores on the EDE-Q, when adjusted for rapid response (by session 8), in a specialist eating

disorders service. Beail et al., (2007) found no differences in treatment responses were observed between patients with intellectual disabilities who received a total of 8, 16 or 24 or more sessions of psychodynamic psychotherapy. To date, there is scarce and inconclusive evidence of a dose-response effect in clinical samples with chronic and/or severe disorders (e.g., psychiatric outpatients), patients with eating disorders, and patients with intellectual disabilities.

## Discussion

### Main findings

With few exceptions, we found replicated and consistent support for the observation of a curvilinear relationship between the number of sessions and the probability of response to treatment, illustrating a trend of diminishing gains with extended treatment. This trend was evident across multiple studies in university counselling centres, outpatient psychotherapy and primary care settings, and across studies of between-subject comparisons and studies that model within-subject trajectories of change. Studies investigating subgroups of patients with differential response trajectories have identified some cases that follow a linear trend of improvement (e.g., see Owen et al., 2015), although such cases are less prevalent (<20%). We also found consistent support for the GEL model assumption that different subgroups of patients show different rates of improvement, suggesting that the duration of treatment is reflective of therapeutic responsiveness. Whilst the GEL and dose-response models may appear contradictory, their implications may not be as mutually exclusive as has been implied in the literature. At an individual level, it is clear that some patients have a rapid response to therapy, whereas others require longer treatments to attain remission of symptoms. However, seen from a population perspective, it is rare to observe cases that attain clinically significant improvements in less than 4 sessions. For example, studies with the briefest dose-response recommendations for psycho-educational interventions in primary care reported than less than

10% of cases overall attained remission of symptoms in treatments with shorter durations than 4 sessions (e.g., Delgadillo et al., 2016). Studies that examine the point at which 75% of cases respond to treatment equally indicate that the probability of response (<25%) to treatment beyond 26 sessions is small (e.g., Howard et al., 1986). Hence, both perspectives are valid: patients with different needs (e.g., levels of severity or impairment) use a differential number of sessions necessary to improve, but usually within predictable boundaries which can be referred to as an optimal dose.

In general, patients with common mental health problems who respond to psychological treatment fall within two distinctive "rapid response" and "gradual response" groups. Rapid responders show signs of improvement by session 4, and gradual responders may need up to 26 sessions to attain reliable and clinically significant improvement. The probability of improvement considerably diminishes with lengthier interventions, and therefore patients that have not shown reliable improvement by session 26 are statistically more likely to be non-responders. These optimal dose parameters were found to vary according to the severity of clinical problems and intensity of treatments. For example, the optimal dose for patients with mild-to-moderate depression and anxiety problems accessing low intensity treatments has been found to be between 4 – 6 sessions; whereas poly-symptomatic clinical samples with varying levels of severity accessing high intensity treatments require between 4 – 26 sessions.

A weekly therapy schedule has been found to accelerate the course of improvement by comparison to a fortnightly treatment schedule (Erekson et al., 2015). This finding fits with meta-analytic evidence that more frequent treatment schedules (e.g., twice per week vs. once per week) are more effective for the treatment of depression (Cuijpers, Huibers, Ebert, Koole, & Andersson, 2013). It is also evident that the necessary dose of treatment varies according to the outcomes of interest; requiring relatively short treatments for 50% response in acute symptoms and longer treatments for improvements in functioning and characterological

problems (Kopta et al., 1994). However, we note that the latter finding is supported by a single study. Further research is required to confirm the replicability and generalisability of findings about the frequency of sessions and change across different outcome domains.

Exceptions to these general optimal dose parameters were found for cases with more chronic and/or severe psychopathology. Some have recommended up to 30 sessions for patients with psychotic disorders (Lincoln et al., 2016). However, there is mixed and inconclusive empirical support for the efficacy of psychotherapy for psychotic disorders in general, with some meta-analyses indicating small effects compared to control groups (Jauhar et al., 2014) and others suggesting no significant effects (e.g., Laws, Darlington, Kondel, McKenna, & Jauhar, 2018). If psychotherapy should exert small effects (e.g., ~ $g$ = 0.33 in Jauhar et al., 2014), the equivalent response rate would be ~17%, and applying the conventional 50% dose-response criterion seems meaningless in this context. Furthermore, no evidence to support the dose-response effect was found in clinical samples of patients with eating disorders (Bell et al., 2016) or those with intellectual disabilities (Beail et al., 2007). Studies in these specific populations converge on the observation that most responders tend to be identified within the first 8 weeks of therapy, and there is little or no evidence that extended therapy will be effective for those who have not shown reliable improvement early on.

In general, there was little evidence to support long-term psychotherapies beyond 30 sessions; and this specific dose-response estimate is for cases with psychotic disorders. The only study that suggested a dose-response effect with an upper boundary of 54 sessions (Asay et al., 2002) was highly atypical, since it analysed cases treated by a single therapist using psychodynamic and integrative therapy. Numerous studies applying psychodynamic and integrative approaches aggregating results across multiple therapists proposed considerably lower optimal dose parameters with an upper boundary of 16 sessions (Erekson et al., 2015; Falkenstrom et al., 2016; Harnett et al., 2010; Kadera et al., 1996; Kopta et al., 1994; Wolgast

et al., 2003). Like the study by Asay et al (2012), some of these studies offered time unlimited therapies (e.g., Anderson & Lambert, 2001) but still observed optimal doses within a relatively short treatment duration when results were aggregated over multiple therapists. Overall, it is important to note that most of the reviewed evidence is from university counselling centres and outpatient psychotherapy clinics for common mental health problems. Many such clinics impose limits on treatment duration, which may constrain the results on the dose-response effect. There is still scarce and inconclusive evidence concerning the relation between treatment duration and outcomes in severe and chronic mental disorders.

**Methodological limitations**

A number of limitations should be considered when interpreting the results of this review. Firstly, studies included in the review did not include randomization or control groups. We cannot therefore assert a cause-effect relationship between therapy dose and clinical outcomes. We can only discern that there is apparently a maximum time period within which remission events are observed (these events may be due to psychotherapy or other factors that promote natural remission). We sought to particularly examine the dose-response effect in routine practice settings because therapies are typically of varying durations and include diverse client groups, thus offering insights that are externally valid to ordinary healthcare populations.

Some studies were excluded from the review as a result of no age range being reported (as per exclusion criteria defined a priori). Many studies demonstrated relatively low quality scores. Only 5 studies were found to adequately adjust for all potential confounding factors and many studies did not report the presenting problems/diagnoses and types of psychotherapies delivered. In particular, little or no information was available in most studies to discern the extent to which concurrent pharmacological treatment may or may not influence the results of dose-response analyses. Many studies (n=18) also failed to adequately follow-up participants and this included identifying any differences between completers and non-

completers of therapy, between censored and non-censored cases in survival analyses or between patients included and excluded in studies when archival data was used. As shown in Table 1, most studies have focused on identifying the point at which 50% of cases respond to treatment, but few have reported the number of sessions required to observe higher (e.g., 75% or 95%) proportions of responders.

Whilst this review has revealed the dose of psychotherapy necessary to enable initial remission of symptoms, the evidence base regarding the longer-term durability of that change is paltry. This may reflect the funding arrangements for routine services; they are commissioned to provide treatment and this does not usually involve longer-term follow-up to examine relapse or recurrence after treatment. In addition, several studies (n=18) were considered to be limited in generalizability on account of all or most of the sample being treated by trainee therapists.

We also note that most reviewed studies conceptualise the dose-response on the basis of symptom reduction. Whilst symptom-reduction as an important target for psychotherapy, we still know relatively little about the potential influence of treatment duration on other outcomes such as prevention of self-harm or suicidality, occupational functioning or impact on the person's wider family and social functioning.

**Recommendations for research**

Future studies could attempt to study the dose-response effect in diagnostically homogeneous clinical samples, or by identifying optimal dose parameters for specific protocol-driven therapies, adjusted by initial symptom severity and demographic characteristics, alongside number of sessions. Introducing specificity in this manner could help to elucidate some inconsistencies in the current evidence base and to develop more fine-grained dose-response expectations that could inform treatment planning (e.g., for different common mental disorders and evidence-based interventions). It is suggested that survival analysis with cox regression,

adjusting for covariates of interest, is a particularly appropriate method to identify optimal doses. Researchers using this method tend to report the estimated number of sessions required for 50% of patients to attain improvement, however it is also important to identify the point at which diminishing gains are observed. We would recommend that the upper boundary could be identified at the 95% percentile using Kaplan-Meier curves, thus indicating the point after which a negligible (<5%) probability of improvement is expected with additional sessions. Furthermore, we argue that the strictest definition of improvement (RCSI, rather than reliable change) is necessary to yield optimal dose recommendations that aim to achieve full remission of symptoms. As previously noted, long-term follow analyses would add considerable value to the dose-response evidence base.

**Implications for policy and practice**

Notwithstanding the heterogeneity of the reviewed studies, discernible patterns of findings in the dose-response literature are evident, with important implications for clinical practice. It is clear that remission of symptoms is rarely observed in extremely brief treatments, and therefore it is recommended that psychological therapy services offer at least 4 sessions as a minimally acceptable dose of treatment. Specifically, in low-intensity guided self-help (GSH) interventions for mild-to-moderate conditions, session 4 is a key point to formally review progress and to address potential obstacles to improvement, since little improvement is typically observed after 6 sessions in this treatment model. Non-responders to GSH after 6 sessions should be referred to more intensive psychotherapies as recommended by clinical guidelines (National Institute for Health and Clinical Excellence, 2011); and extending the duration of GSH beyond 6 sessions should be limited to cases that show signs of reliable improvement. In more conventional psychotherapies including moderate-to-severe cases, session 8 tends to mark the point at which the bulk of symptomatic improvement is observed, and therefore a formal review of progress is advisable. At this review point, potential

difficulties in therapy or obstacles to improvement could be assessed for cases that have not yet shown signs of reliable improvement. Adjusting therapy to address potential obstacles to improvement (e.g., motivation to change, non-compliance with therapeutic strategies, therapeutic alliance deficits, social support deficits, etc.) may be necessary for a trial period (e.g., up to 26 sessions to allow the identification of the majority of treatment responders). The evidence for longer-term treatments is still inconclusive, and the quality of reviewed studies was moderate, so future research is necessary to provide more decisive guidelines on the circumstances when longer-term psychotherapy may be indicated. In practice, decisions about extending treatment should not only be based on the probability of symptomatic improvements, but also considerations related to potential risks to self or to others.

**Disclosure of interest:** The authors report no conflicts of interest.

**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

Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of Change in Naturalistic Psychotherapy: Contrasting Dose-Effect and Good-Enough Level Models of Change. *Journal of Consulting and Clinical Psychology*, *77*(2), 203–211. https://doi.org/10.1037/a0015235

Barkham, M., Connell, J., Stiles, W. B., Miles, J. N., Margison, F., Evans, C., & Mellor-Clark, J. (2006). Dose-effect relations and responsive regulation of treatment duration: The good enough level. *Journal of Consulting and Clinical Psychology, 74*(1), 160-167.

Barkham, M., Gilbert, N., Connell, J., Marshall, C., & Twigg, E. (2005). Suitability and utility of the CORE-OM and CORE-A for assessing severity of presentig problems in psychological therapy services based in primary and secondary care. *British Journal of Psychiatry*, *186*, 239–246. https://doi.org/10.1192/bjp.186.3.239

Barkham, M., Rees, A., Stiles, W. B., Shapiro, D. A., Hardy, G. E., & Reynolds, S. (1996). Dose-effect relations in time-limited psychotherapy for depression. *Journal of Consulting and Clinical Psychology*, *64*(5), 927–935. https://doi.org/10.1037/0022-006X.64.5.927

Beail, N., Kellett, S., Newman, D. W., & Warden, S. (2007). The Dose – Effect Relationship in Psychodynamic Psychotherapy with People with Intellectual Disabilities. *Journal of Applied Research in Intellectual Disabilit*ies, *20,* 448–454.

Bell, C., Waller, G., Shafran, R., & Delgadillo, J. (2016). Is there an optimal length of

psychological treatment for eating disorder pathology? *International Journal of Eating Disorders*, *50*(6), 687–692. https://doi.org/10.1002/eat.22660

Bennett-Levy, J., Richards, D.A., Farrand, P (2010). Low intensity CBT interventions: a revolution in mental health care; in Bennett-Levy J, Richards DA, Farrand P, Christensen H, Griffiths KM, Kavanagh DJ, et al. (eds): Oxford Guide to Low Intensity CBT Interventions. Oxford, Oxford University Press, 2010

Callahan, J. L., & Hynan, M. T. (2005). Models of Psychotherapy Outcome: Are They Applicable in Training Clinics? *Psychological Services*, *2*(1), 65–69. https://doi.org/10.1037/1541-1559.2.1.65

Carr, M. M., Saules, K. K., Koch, E. I., & Waltz, T. J. (2017). Testing the dose-response curve in a training clinic setting: Use of client pretreatment factors to minimize bias in estimates. *Training and Education in Professional Psychology*, *11*(1), 26–32. https://doi.org/10.1037/tep0000135

Cuijpers, P., Huibers, M., Ebert, D. D., Koole, S. L., & Andersson, G. (2013). How much psychotherapy is needed to treat depression? A metaregression analysis. *Journal of Affective Disorders, 149*(1-3), 1-13. https://doi.org/10.1016/j.jad.2013.02.030

Delgadillo, J., Kellett, S., Ali, S., McMillan, D., Barkham, M., Saxon, D., Donohoe, G., Stonebank, H., Mullaney, S., Eschoe, P., Thwaites, R., & 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

Draper, M. R., Jennings, J., Baron, A., Erdur, O., & Shankar, L. (2002). Time-Limited

Counseling Outcome in a Nationwide College Counseling Center Sample. *Journal of College Counseling*, *5*(1), 26–38. https://doi.org/10.1002/j.2161-1882.2002.tb00204.x

Erekson, D. M., Lambert, M. J., & Eggett, D. L. (2015). The relationship between session frequency and psychotherapy outcome in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, *83*(6), 1097–1107. https://doi.org/10.1037/a0039774

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, *180*, 51–60. https://doi.org/10.1192/bjp.180.1.51

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorder psychopathology: Interview or self-report questionnaire. *International Journal of Eating Disorders*, *16*(4), 363–370.

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. (2003). An evaluation of the dose-response relationship in naturalistic treatment settings using survival analysis. *Mental Health Services Research*, *5*(1), 1–12. https://doi.org/10.1023/A:1021751307358

Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, *9*(3), 329–343. Retrieved from http://onlinelibrary.wiley.com/o/cochrane/cldare/articles/DARE-12002006550/frame.html

Harnett, P., O'Donovan, A., & Lambert, M. J. (2010). The dose response relationship in psychotherapy: Implications for social policy. *Clinical Psychologist*, *14*(2), 39–44.

https://doi.org/10.1080/13284207.2010.500309

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 defining 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

Jauhar, S., McKenna, P. J., Radua, J., Fung, E., Salvador, R., & Laws, K. R. (2014). Cognitive–behavioural therapy for the symptoms of schizophrenia: systematic review and meta-analysis with examination of potential bias. *British Journal of Psychiatry, 204*(1), 20-29. https://doi.org/10.1192/bjp.bp.112.116285

Kadera, S. W., Lambert, M. J., & Andrews, a a. (1996). How much therapy is really enough? A session-by-session analysis of the psychotherapy dose-effect relationship. *The Journal of Psychotherapy Practice and Research*, *5*, 132–151. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22700273%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3330412

Kopta, S. M., Howard, K. I., Lowry, J. L., & Beutler, L. E. (1994). Patterns of symptomatic recovery in psychotherapy. *Journal of Consulting and Clinical Psychology*, *62*(5), 1009–1016.

Kopta, S. M., & Lowry, J. L. (2002). Psychometric evaluation of the Behavioral Health Questionnaire-20: A brief instrument for assessing global mental health and the three phases of psychotherapy outcome. *Psychotherapy Research*, *12*(4), 413–426. https://doi.org/10.1093/ptr/12.4.413

Kopta, S. M., Petrik, M. L., Saunders, S. M., Mond, M., Hirsch, G., Kadison, R., & Raymond, D. (2014). The utility of an efficient outcomes assessment system at university counseling

centers. *Journal of College Student Psychotherapy*, *28*(2), 97–116. https://doi.org/10.1080/87568225.2014.883876

Lambert, M. J., Burlingaume, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology & Psychotherapy*, *3*(4), 249–258.

Laws, K. R., Darlington, N., Kondel, T. K., McKenna, P. J., & Jauhar, S. (2018). Cognitive Behavioural Therapy for schizophrenia-outcomes for functioning, distress and quality of life: a meta-analysis. *BMC Psychology, 6*(1), 32. https://doi.org/10.1186/s40359-018-0243-2

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.

Lutz, W., Schürch, E., Stulz, N., Böhnke, J. R., Schöttke, H., Rogner, J., & Wiedl, K. H. (2009). Entwicklung und psychometrische Kennwerte des Fragebogens zur Evaluation von Psychotherapieverläufen (FEP). *Diagnostica*, *55*(2), 106–116. https://doi.org/10.1026/0012-1924.55.2.106

National Institute for Health and Clinical Excellence (2011). *Common mental health disorders: identification and pathways to care*. National Collaborating Centre for Mental Health, London. Available at: http://guidance.nice.org.uk/CG123/NICEGuidance/pdf/English.

Owen, J., Adelson, J., Budge, S., Wampold, B., Kopta, M., Minami, T., & Miller, S. (2015). Trajectories of change in psychotherapy. *Journal of Clinical Psychology*, *71*(9), 817–827. https://doi.org/10.1002/jclp.22191

Owen, J. J., Adelson, J., Budge, S., Kopta, S. M., & Reese, R. J. (2016). Good-enough level and dose-effect models: Variation among outcomes and therapists. *Psychotherapy*

*Research*, *26*(1), 22–30. https://doi.org/10.1080/10503307.2014.966346

Reese, R. J., Toland, M. D., & Hopkins, N. B. (2011). Replicating and extending the good-enough level model of change: Considering session frequency. *Psychotherapy Research*, *21*(5), 608–619. https://doi.org/10.1080/10503307.2011.598580

Rubel, J., Lutz, W., Kopta, S. M., Köck, K., Minami, T., Zimmermann, D., & Saunders, S. M. (2015). Defining early positive response to psychotherapy: An empirical comparison between clinically significant change criteria and growth mixture modeling. *Psychological Assessment, 27*(2), 478. DOI: 10.1037/pas0000060

Snell, M. N., Mallinckrodt, B., Hill, R. D., & Lambert, M. J. (2001). Predicting counseling center clients' response to counseling: A 1-year follow-up. *Journal of Counseling Psychology*, *48*(4), 463–473. https://doi.org/10.1037/0022-0167.48.4.463

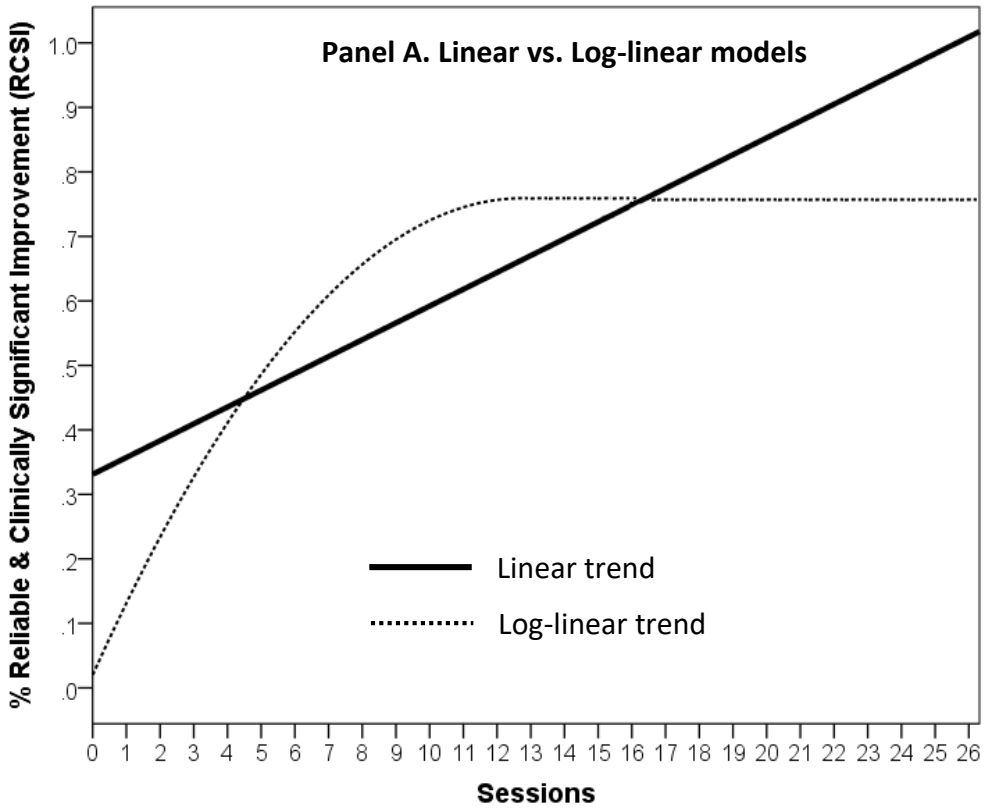
Stiles, W. B., Barkham, M., Connell, J., & Mellor-Clark, J. (2008). Responsive Regulation of Treatment Duration in Routine Practice in United Kingdom Primary Care Settings: Replication in a Larger Sample. *Journal of Consulting and Clinical Psychology*, *76*(2), 298–305. https://doi.org/10.1037/0022-006X.76.2.298

Stiles, W. B., Barkham, M., & Wheeler, S. (2015). Duration of psychological therapy: Relation to recovery and improvement rates in UK routine practice. *British Journal of Psychiatry*, *207*(2), 115–122. https://doi.org/10.1192/bjp.bp.114.145565

Stulz, N., Lutz, W., Kopta, S. M., Minami, T., & Saunders, S. M. (2013). Dose-effect relationship in routine outpatient psychotherapy: Does treatment duration matter? *Journal of Counseling Psychology*, *60*(4), 593–600. https://doi.org/10.1037/a0033589

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

**Figure 1. Linear versus non-linear change models**



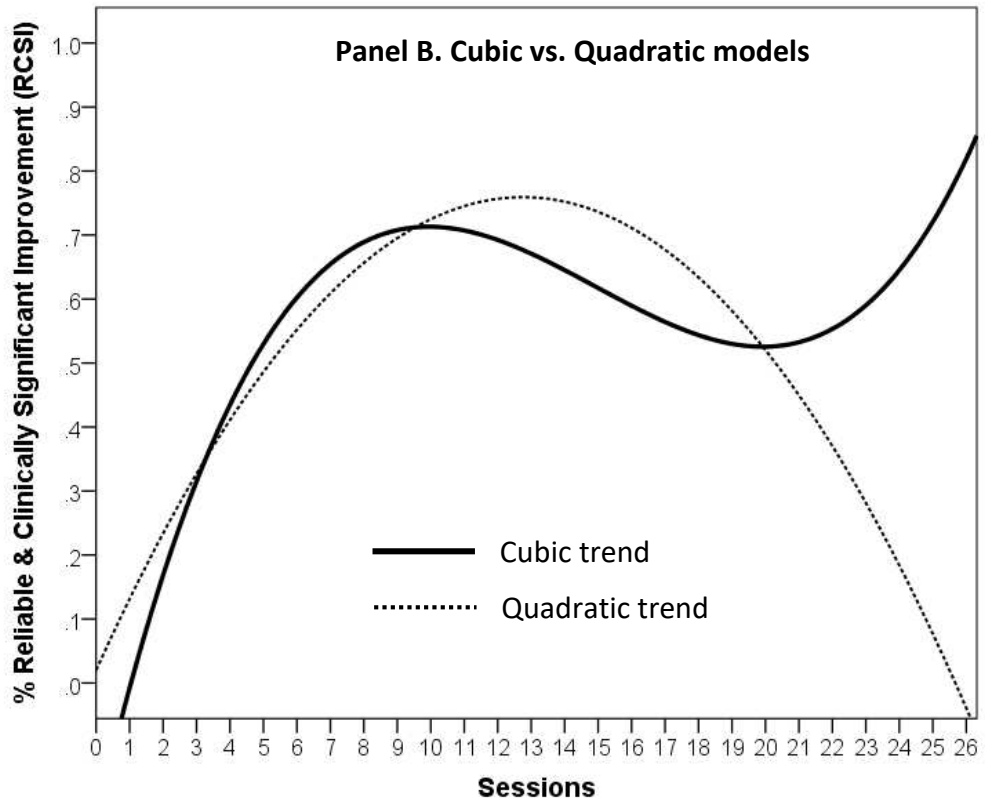
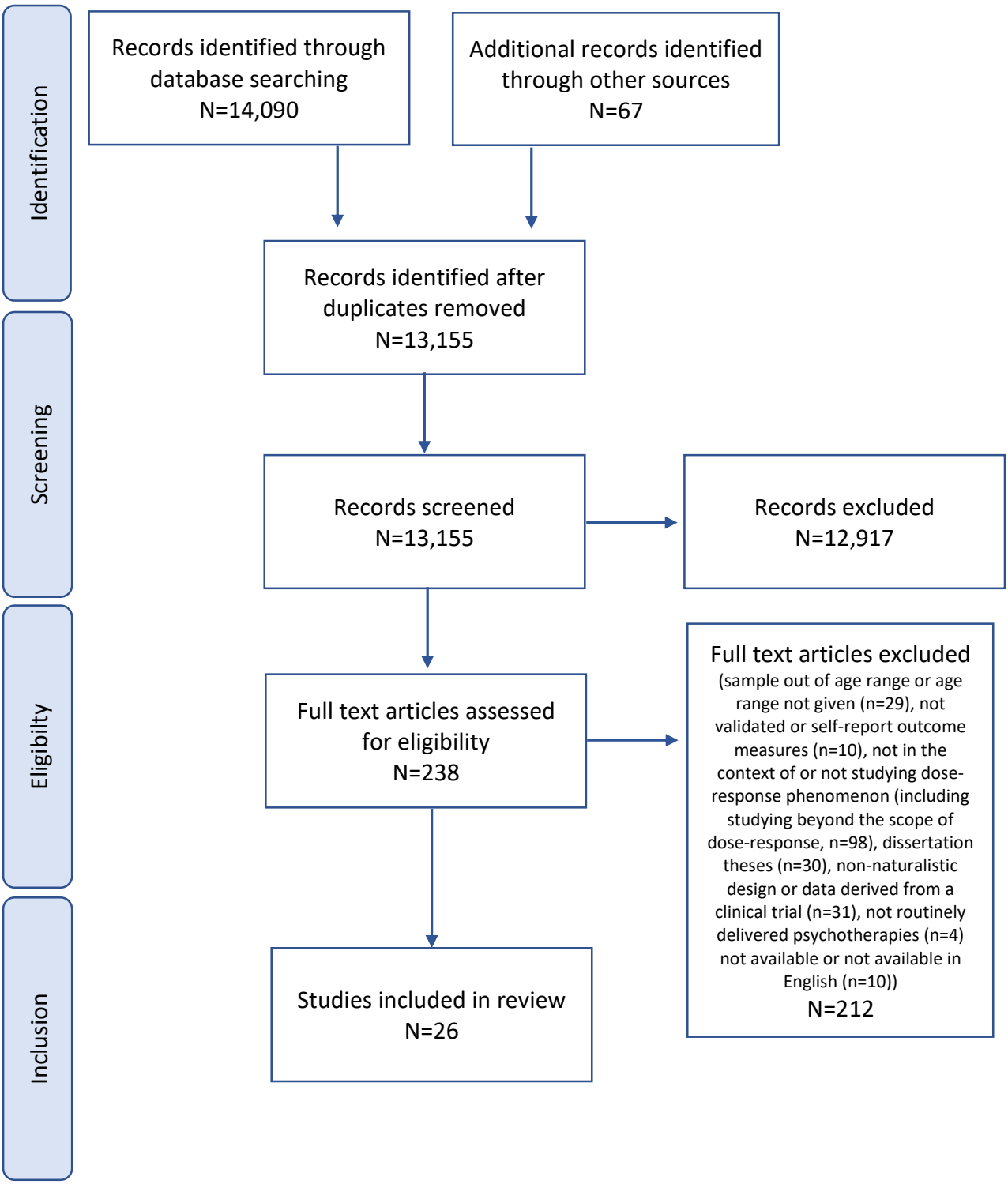



**Figure 2. PRISMA Diagram**



## Table 1. Characteristics of studies included in the review

| Article | Sample size | Patient demographics | Symptoms and diagnoses | Interventions | Setting | Outcome measures | How was dose-response studied? | Findings on dose-response | Optimal dose | Quality Score |
|---|---|---|---|---|---|---|---|---|---|---|
| Anderson & Lambert (2001) | N= 75 | 18-52 (M=30)<br>65.3% female, 34.7% male | Mixed | Mixture of low-cost, time unlimited services. Therapist orientations described as eclectic. | University training clinic, USA. | OQ-45 | Kaplan Meier Survival Analysis | 8 sessions for 50% to RC<br>11 sessions for 50% to RCSI | 8-11[1] | 9 |
| Asay et al., (2002) | N= 29 (adult sample only reported here) | Adults 19+.<br>No demographics provided. | Mixed | Primarily psychodynamic, integrated with other approaches. | Private practice, USA. | OQ-45 | Kaplan Meier Survival Analysis with Cox regression | 42 sessions for 50% to RC<br>54 sessions for 50% to RCSI | 42-54[1] | 7 |
| Baldwin et al., (2009) | N=4676 | 17-60 (M=22.3)<br>62% female. | Mixed | Individual therapy | Archival dataset from University counselling service, USA. | OQ-45 | Multilevel modelling<br><br>Logistic regression | Shape of change best modelled by a cubic function of sessions.<br><br>Dose effect model: in early therapy there is a rapid rate of change and this flattens between 8 and 10 sessions.<br><br>However better fit in stratified model in favour of GEL model.<br><br>Logistic regression suggests increase in probability of achieving RCSI until 8 sessions. | 8-10[2] | 10 |
| Beail et al., (2007) | N= 20 | 17-48 (M=29.3)<br>15% female, 85% male. | All patients presented with intellectual disability. Therapy was for a mixture of psychological symptoms. | Psychodynamic psychotherapy | Service for people with intellectual disabilities, UK | SCL-90-R, IIP-32, Rosenberg Self-Esteem Inventory | ANOVA | No differences in treatment responses between 8, 16 and 24+ sessions.<br><br>8 sessions is sufficient to observe no further gains.<br><br>Study suggests this supports the dose-response model. | 8[2] | 6 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Bell et al., (2016) | N=164 | 18-60 (M=30.13)<br>95.1% female | Eating disorders | Integrative psychotherapy (41.5%), cognitive behavioural therapy (18.3%), person centered counselling (12.2%), cognitive analytic therapy (12.2%), gestalt therapy (11.6%) and EMDR (4.3%) | Outpatient psychological care service for eating disorders, UK. | EDE-Q | MANOVA | No association found between treatment duration and outcome on EDE-Q when adjusted for rapid response early in treatment. | No association | 11 |
| Callahan & Hynan, (2005) | N=61 | 18-55 (M=29.6)<br>52% female | Mixed | Not stated | University training centre, USA. | OQ-45 | Percentages of patients recovering (includes both RC/RCSI in the same outcome category). | **Responsiveness appears to diminish around 24 sessions<br><br>In 26 sessions, 31% recovered.<br><br>Did not reach 50% recovered within the 52 session duration examined. | 24[2] | 7 |
| Carr et al., (2017) | N=132 | 18+ years (M=30)<br>60.3% female | Mixed | Individual therapy only | Training clinic, USA. | OQ-45 | Kaplan Meier Survival Analysis with number of PHQ factors entered as a factor. | 7 sessions for 50% to RC<br>19 sessions for 50% to RCSI<br><br>This differed by number of PHQ diagnoses. | 7-19[1] | 7 |
| Delgadillo et al., (2016) | N=4451 | 16-89 (M=42.94)<br>63.1% female | Anxiety and depression related symptoms | Low intensity stress control group therapy | 5 IAPT centres, UK. | PHQ-9, GAD-7 | Percentages of patients achieving RCSI | Greatest gains first observed at 4 sessions with diminishing gains beyond 6 sessions | 4-6[2] | 11 |
| Delgadillo et al., (2014) | N=1850 | 16-87 (M=37.85)<br>65% female, 35% male | Anxiety and depression related symptoms | Low intensity therapies delivered in routine care on the IAPT programme. | IAPT service, UK. | PHQ-9, GAD-7 | Percentages of patients achieving RCSI | Minimum of 4 sessions for 50% RCSI<br><br>Diminishing gains beyond 6 sessions | 4-6[2] | 10 |
| Draper et al., (2002) | N=1698 | *Mean age=23.1 (SD=5.6)<br>66.5% female<br>33.5% male | Not stated. | 95% of centres included used a brief (10 session) model of therapy. | 42 universities and 9 private institutions (all higher education facilities), USA. | OQ-45 | Percentage of patients improved (includes both RCSI and RC in a combined category) | Peaks in improvement rates at sessions 4, 6 and 10.<br><br>Minimum number of sessions suggested is 4, however 49.5% improved at 6 sessions and most improvement observed at session 10 (54.5% improved). | 4-10[3] | 5 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Erekson et al., (2015) | N= 21488 | *Mean age= 22.5 54.9% female, 39.8% male, 5.3% unspecified. | Mixed | Individual therapy only. Types of therapy delivered guided by therapist orientations, but included cognitive–behavioural, psychodynamic, client-centred, existential, systems, and integrative modalities | University counselling centre, USA. | OQ-45 | Multilevel modelling<br><br>Multilevel cox regression<br><br>Multilevel logistic regression | Cubic modelling of sessions provides best fit to the data.<br><br>Total dose significantly improves model fit, supporting GEL over dose-effect.<br><br>Faster recovery is achieved in once weekly rather than fortnightly therapy. RCSI is equally likely in both weekly and fortnightly groups.<br><br>Increased session efficacy if dosage received once a week, with 50% achieving RC in 8 sessions. | 8[1] | 8 |
| Falkenstrom et al., (2016) | Primary care sample: N=640<br>Psychiatric outpatient sample: N=284 | Primary care sample: 16-88 (m=36.8), 55.9% female , 18% male, remaining data missing<br>Psychiatric outpatient sample:16-64 (M=32.6), 36.1% male, 33.3% female, remaining data missing | Mixed. Psychiatric sample presented with more severe symptoms. | Primary care sample: Individual therapies including supportive (30%), psychodynamic (24%), CBT (18%), crisis intervention (15%), cognitive (15%), behavioural (9%), relational (9%), existential (7%), systemic (7%), and interpersonal (6%)<br>Psychiatric outpatient sample: CBT (48%), psychodynamic (44%), supportive therapy (37%), relapse prevention (29%), and Motivational Interviewing (28%). As in the primary care sample, many therapists reported more than one orientation for a single therapy. | Community based primary care setting and adult psychiatric outpatient clinic, Sweden. | CORE-OM | Multilevel modelling | Cubic modelling of sessions provides best fit to the data in primary care sample only.<br><br>From the aggregate model, additional benefits observed between 8-13 sessions in primary care sample. Linear improvement observed in psychiatric sample.<br><br>Faster rates of recovery observed in patients attending fewer sessions, supporting GEL model.<br><br>Slower overall recovery rates in psychiatric sample.<br><br>Primary care sample improved more than psychiatric sample. | 8-13[4] | 9 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Harnett et al., (2010) | N=125 | 18-65 (M= 34.5 )<br>65% female, 35% male. | Mixed | Therapist orientations varied but included cognitive behavioural, interpersonal, and psychodynamic. | 2 university training clinics, Australia. | OQ-45 | Kaplan-Meier Survival Analysis | 10 sessions for 50.6% to RC<br>14 sessions for 50% to RCSI | 10-14[1] | 7 |
| Kadera et al., (1996) | N=64 | All adults<br>(M=28-33)<br>Males (31%), females (69%) | Mixed | Cognitive behavioural, humanistic-existential, psychodynamic interpersonal, integrative-eclectic. | University training facility, USA. | OQ-45 | Percentage of patients who reach RCSI at different dosages.<br><br>Recovery tables, based on survival analysis | Of 21 patients who did reach RCSI:<br>8 sessions for 43% RCSI<br>13 sessions for 76% RCSI<br><br>**Observation of the dose-effect curve presented in the study suggests around 10 sessions for 50% RCSI<br><br>Estimates for the dysfunctional sample based on recovery tables:<br>13 sessions for 44% RCSI<br>16 sessions for 50% RCSI | 10-16[1] | 8 |
| Kopta et al., (1994) | N=854 | 18+ (M=28-39)<br>68-74% female across the 5 sites | Mixed | Therapist orientations included psychodynamic, eclectic, and other e.g. cognitive-behavioural, client centred. | 5 mental health centres delivering psychotherapy, USA. | SCL-90-R | Probit analysis | Median effective dose for RCSI:<br>5 sessions for acute symptoms<br>14 sessions for chronic symptoms<br>>52 sessions for characterological symptoms<br>11 sessions for most common symptoms | 11[1] | 10 |
| Kopta et al., (2014) | N=13,803 | 18+<br>62.5% female, 35.3% male | Not stated | Therapists varied in orientation and approach. | 23 university counselling centres, USA. | BHM-20 with Suicide Monitoring Scale | Percentage of patients who reach RC as different dosages.<br><br>Chi square analyses<br><br>ANOVA | Change scores significantly greater at 7-10 sessions, with diminishing gains beyond this point.<br><br>Percentage RC significantly greater at 7-10 sessions, with no significantly greater gains beyond this. | 7-10[2] | 6 |
| Lincoln et al., (2016) | N=58 | 18-65 (M=35.67)<br>40% female, 60% male | All patients presented with psychosis | 45 session model of CBT for psychosis | Outpatient clinic of a psychology department, Germany. | CAPE, SCL-27plus, CHOICE | Linear Multilevel modelling<br><br>Multilevel logistic regression | First significant reduction in scores at 15 sessions, with minimal gains beyond 25.<br><br>Corroborated on the whole by CS rates measured on 2 outcome scales. | 25-30[2] | 11 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Diminishing gains between 20 and 30 sessions. | | |
| Owen et al., (2016) | N= 13,664 | *Adults<br>65.4% female, 31.2% male, 3.4% not reported. | Not collected. | Not stated. | 46/47 University counselling centres, 1 other, USA. | BHM-20 | Multilevel modelling | Log-linear modelling of sessions provides best fit to the data, except for the life-functioning subscale, which was best represented by a quadratic function.<br><br>Models including fixed effect of total dose and interactions consistently out-perform dose-effect aggregate models. Supports GEL model. | n/a | 6 |
| Owen et al (2015) | N= 10854 | *Adults<br>61.8% female, 31.5% male, 6.7% not reported. | Not collected. | Not stated. | 46 University counselling centres and 1 community outpatient setting, USA. | BHM-20 | Multilevel growth mixture modelling | 3 latent classes of patients with different change trajectories: worse before better, slow and steady and early and late responders.<br><br>The 'early and late change' class represented over 75% of the sample. Within this group, patients needed to remain in therapy for a minimum of 11 sessions to benefit from the second period of improvement. | 11+[4] | 6 |
| Reese et al., (2011) | N=1207 | 17.64-63.69 (M=23.72)<br>69.7% female | Mixed | Cognitive behavioural, solution focused, psychodynamic, interpersonal process, narrative. Mostly a combination. | University counselling centre, USA. | OQ-45 | Multilevel modelling | Linear modelling of sessions provides best fit to the data based on visual inspection of the data.<br><br>Including total dose of therapy in the model significantly improved model fit. Supports GEL model.<br><br>Including session frequency as a fixed effect significantly improved model fit. Weekly better than fortnightly. | n/a | 8 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sembill et al., (2017) | N=351 | 20-75 (M=41.1)<br>65.2% female | Mixed | CBT or Psychodynamic therapy | University outpatient clinic, Germany | FEP | Multilevel modelling | Log-linear modelling of sessions provides best fit to the data for the global FEP score. | n/a | 8 |
| Snell et al., (2001) | N=158 | 18-60, (M=25.96)<br>63% Female, 37% male. | Not stated | 12 session therapy limit. | University counselling centre, USA. | CASPER-13F, OQ-45 | Kaplan-Meier Survival Analysis with Cox regression<br><br>Percentage of clients to reach RC and RCSI compared across dosage categories with chi square analyses | 14 sessions for 50% to RC at 10 month follow up<br>16 sessions for 50% to RCSI at 10 month follow up<br><br>Cox regression shows no significant effect of pre-test scores on CS change but smaller dosage required to achieve RC if higher pre-test scores are observed.<br><br>Significantly less patients attending 2-7 sessions achieved RCSI than patients attending 1 or 8+ sessions. No differences observed for RC in treatment lengths. | 14-16[1]<br><br>8+[4] | 8 |
| Stiles et al., (2008) | N=9703 | 16-99(M=40.9)<br>72.4% female | Mixed | Integrative (40.4%), person-centred (37.0%), structured/brief (31.4%), cognitive behavioural (26.4%), supportive (16.8%), psychodynamic (16.1%). 52.6% received a mix. | 32 NHS primary care services, UK. | CORE-OM | Percentage RCSI rates correlated with total number of sessions attended. | RCSI rates negatively correlated with total number of sessions (r=-.75).<br><br>RC was not significantly correlated with total number of sessions.<br><br>Supports GEL and responsive regulation model. | n/a | 7 |
| Stiles et al., (2015) | N=26430 | 16-95 (m=38.6)<br>69.3% female | Mixed | Integrative (41.2%), person-centred (36.4%), psychodynamic (22.8%), cognitive behavioural (14.9%), structured/brief (14.6%),supportive (14%). 41.6% more than 1 type of therapy. | 50 services, UK.<br>(6 primary care services, 8 secondary, 2 tertiary, 10 university 14 voluntary 8 workplace and 2 private practices). | CORE-OM | Percentage RCSI rates correlated with total number of sessions attended. | RCSI rates were negatively correlated with total number of sessions (r=-.58)<br><br>RC rates were negatively correlated with total number of sessions (r=-.40).<br><br>Pattern observed in multiple sectors but not all. | n/a | 9 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Supports GEL and responsive regulation model. | | |
| Stulz et al., (2013) | N=6375 | 18+<br>64% female, | Not stated | Variety of theoretical orientations. | 26 centres including 20 college counselling centres (accounting for 94% of clients), 4 primary care (5.8%), 2 private hospitals (0.2%). | BHM | Multilevel latent growth curve modelling.<br><br>Percentage RCSI rates correlated with total number of sessions attended. | Log-linear modelling of sessions provides best fit to the data.<br><br>Rate of change is faster in shorter therapies. This corroborates GEL but also dose-effect as negatively accelerated regardless of treatment duration.<br><br>Total length of therapy was positively correlated with RCSI rates. | n/a | 9 |
| Wolgast et al., (2003) | N=788 | 18.1=63.6 (M=23.2), 66.6% Female, 33.3% male, | Mixed | Individual. Only 20 of 43 staff listed orientations which included: psychoanalytic, psychodynamic, integrative (cognitive behavioural, psychodynamic and interpersonal), eclectic. | University counselling service, USA. | OQ-45 | Kaplan-Meier Survival Analysis | 10 sessions for 51% RC<br>14 sessions for 51% RCSI<br><br>**These session numbers also coincide with minimal gains beyond these points, from observation of survival analysis tables. | 10-14[3] | 8 |

*Notes*:
* Indicates adult sample has been assumed
** Observation interpreted by the author from a graph presented in the study
RC, statistically reliable change; CS, clinically significant change based on a caseness cut-off; RCSI, reliable and clinically significant change; GEL, Good Enough Level model; OQ-45, Outcome Questionnaire- 45; SCL-90-R, Symptom Check-list-90 Revised; IIP-32, Inventory of Interpersonal Problems-32; ANOVA, Analysis of Variance; EDE-Q, Eating Disorders Examination Questionnaire; EMDR, Eye Movement and Desensitisation Reprocessing; MANOVA, Multivariate Analysis of Variance; PHQ-9, Patient Health Questionnaire-9; GAD-7, Generalised Anxiety Disorder-7; IAPT, Improving Access to Psychological Therapies programme; CORE-OM, Clinical Outcomes in Routine Evaluation- Outcome Measure; CBT, Cognitive Behavioural Therapy; BHM-20, Behavioural Health Measure-20; CAPE, Community Assessment of Psychic Symptoms; SCL-27plus, Symptom Checklist- 27 plus; CHOICE, CHoice of Outcome In CBT for psychoses; FEP, German Questionnaire for the Evaluation of Psychotherapeutic Change Processes; CASPER-13F, Computerized Assessment System for Psychotherapy Evaluation and Research- follow up, NHS, National Health Service.
[1]based on 50% RCSI or RC
[2]based on diminishing gains
[3]based on a combination of 50% RC or RCSI and diminishing gains
[4]based on a minimum to observe additional gains after a plateau

**Table 2. Summary of outcomes of interest in the included studies.**

| Outcome event or measures of interest | N | Studies |
|---|---|---|
| RCSI at first observed point, maintained until termination | 1 | Kadera et al., (1996) |
| RC, RCSI at first observed point, maintained until termination | 4 | Anderson & Lambert, (2001)<br>Asay et al., (2002)<br>Harnett et al., (2010)<br>Wolgast et al., (2003) |
| RCSI at termination | 5 | Baldwin et al., (2009)<br>Delgadillo et al., (2014)<br>Delgadillo et al., (2016)<br>Stiles et al., (2008)<br>Stiles et al., (2015) |
| RC, RCSI at termination | 1 | Carr et al., (2017) |
| RC, RCSI at 10 month follow-up | 1 | Snell et al., (2001) |
| RC and RCSI combined category | 3 | Callahan & Hynan, (2005)<br>Draper et al., (2002)<br>Kopta et al., (2014) |
| RCSI first observation | 1 | Kopta et al., (1994) |
| RC, RCSI first observation | 1 | Erekson et al., (2015) |
| CS at termination | 1 | Lincoln et al., (2016) |
| Pre-post therapy change scores | 3 | Beail et al., (2007)<br>Bell et al., (2016)<br>Kopta et al., (2014)<br>Stulz et al., (2013) |
| Session-by-session change scores | 7 | Baldwin et al., (2009)<br>Erekson et al., (2015)<br>Falkenstrom et al., (2016)<br>Owen et al., (2016)<br>Owen et al., (2015)<br>Reese et al., (2011)<br>Sembill et al., (2017) |

*Notes*: RC, statistically reliable change; CS, clinically significant change based on a caseness cut-off; RCSI, reliable and clinically significant improvement.