Article:
Beard, J. and Delgadillo, J. (2019) Early response to psychological therapy as a predictor of depression and anxiety treatment outcomes: A systematic review and meta-analysis. Depression and Anxiety, 36 (9). pp. 866-878. ISSN: 1091-4269

https://doi.org/10.1002/da.22931

Citation: Beard, J.I.L., & Delgadillo, J. (*in press*). Early response to psychological therapy as a predictor of depression and anxiety treatment outcomes: A systematic review and meta-analysis. *Depression and Anxiety*. doi: 10.1002/da.22931

Early response to psychological therapy as a predictor of depression and anxiety treatment outcomes: A systematic review and meta-analysis

Jessica I.L. Beard[1] and Jaime Delgadillo[2*]

[1] Department of Psychology, University of Sheffield, United Kingdom

[2] Clinical Psychology Unit, Department of Psychology, University of Sheffield, United Kingdom

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

ABSTRACT

Background: Previous studies indicate that early symptomatic improvement, typically observed during the first 4 weeks of psychological therapy, is associated with positive treatment outcomes for a range of mental health problems. However, the replicability, statistical significance and magnitude of this association remains unclear.

Aim: The present study reviewed the literature on early response to psychological interventions for adults with depression and anxiety symptoms.

Methods: A systematic review and random effects meta-analysis was conducted, including studies found in Medline, PsychINFO, SCOPUS, Web of Science, and through reference lists and reverse citations.

Results: Twenty-five eligible studies including 11091 patients measured early response and examined associations with post-treatment outcomes. It was possible to extract and/or calculate effect size data from 15 studies to conduct a meta-analysis. A large pooled effect size (g = 0.87 [95 % CI: 0.63, 1.10] $p < .0001$) indicated that early responders had significantly better post-treatment outcomes compared to cases without early response, and this effect was larger in anxiety (g = 1.37) compared to depression (g = 0.76) measures. Most studies were of good quality and there was no evidence of publication bias. The main limitations concerned insufficient statistical reporting in some studies, which meant that the meta-analysis only included 60% of reviewed studies, and it was not possible to examine effect sizes according to different outcome questionnaires.

Conclusions: There is robust and replicated evidence that early response to therapy is a reliable prognostic indicator for depression and anxiety treatment outcomes.

Key words: early response; psychological therapy; depression; anxiety

1. INTRODUCTION

Symptomatic improvements that occur during the initial sessions of psychological therapy, and which are of a statistically and/or clinically significant magnitude, have been termed *early response.* As early as the 1980s, psychotherapy researchers have observed that the bulk of symptomatic improvement occurs within the first month of therapy, with 60% of total improvement in cognitive behavioural therapy for depression occurring by week four (Rush, Kovacs, Beck, Weissenburger, & Hollon, 1981). Since then, numerous other studies have reported early response patterns across different mental health problems and psychological therapies (Bell, Waller, Shafran, & Delgadillo, 2017; Bradford et al., 2011; Delgadillo et al., 2014; Doyle, Grange, Loeb, Doyle & Crosby, 2009; Gois et al., 2014; Hunnicutt-Ferguson, Hoxha & Gollan, 2012; Rubel et al., 2015).

Some studies have found that early symptomatic response is associated with better post-treatment outcomes (e.g., Crits-Christoph et al., 2001; Grilo, Masheb & Wilson, 2006; Lutz et al., 2017; Stauffer et al., 2011; van Calker et al., 2009). However, the magnitude of this association varies considerably across studies and clinical samples. For example, in samples of patients accessing guided self-help for depression, Tadić et al. (2010) reported an odds ratio of 1.33 whereas Delgadillo et al. (2014) reported an odds ratio of 12.60. Furthermore, others have found mixed results. For example, Arnow et al. (2007) found no significant associations between early response and treatment dropout. Başoğlu et al. (1994) found that early response predicted some post-treatment outcomes (panic attacks) but not others (anticipatory avoidance). In addition, there is still no agreed-upon time scale of when early response occurs by in psychological therapy (Haas, Hill, Lambert & Morrell, 2002; Rubel et al., 2015). For example, some studies have shown that this can happen by week two or three, whilst others have examined this by week eight or later (Arnow et al., 2007; Bell et al. 2017; Rubel et al., 2015; Tadić et al., 2010). Also, different psychological therapies have

varying numbers of sessions within the first month of treatment. For example, Tang and DeRubeis (1999a) state that there are not always four sessions within the first month of cognitive behavioural therapy (CBT); instead, in treatment lasting around 12 to 20 weeks, 40 to 60% of sessions occur within the first month. Studies assessing early response are also heterogeneous, since they use different symptom measures and baseline levels of distress differ among study samples.

Therefore, whilst it appears to be that people with early symptomatic improvements tend to have a better prognosis, the replicability, statistical significance and magnitude of the association between early response and post-treatment outcomes remains unclear. The current systematic review aimed to synthesise what is known about the association between early response and psychological treatment outcomes in adult patients seeking treatment for depression and anxiety.

2. METHODS

2.1. Protocol and Registration

The review protocol was prospectively registered in the PROSPERO database (http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018089123).

2.2. Search Strategy and Study Selection

Table 1 specifies the inclusion and exclusion criteria that guided this review. Four databases were searched in order to search for titles, abstracts and key words (Appendix A) on February 13, 2018: Medline, PsychINFO, SCOPUS and Web of Science. The search was limited to peer reviewed articles or book chapters published in English. No date restrictions were applied. A full list of excluded studies and reasons for exclusion are listed in Appendix B.

The search also included the terms “sudden gains”, which can be defined as sudden and statistically significant symptomatic improvements that occur between two consecutive sessions during therapy (Tang & DeRubeis, 1999b). Although sudden gains are conceptually different to early response (which can be gradual and over more than two sessions), sudden gains that occur early in treatment can overlap with and contribute to an early response pattern. Therefore, studies where the median sudden gain session occurred within the first four weeks of therapy were also included; consistent with evidence that most sudden gains occur within the first month (e.g., Tang & DeRubeis, 1999a). Titles and abstracts were screened by the first author, followed by a full-text review. Of the studies eligible for the review, reverse-citations and reference list searches were conducted by hand to identify further eligible studies. A PRISMA diagram summarising the study selection process is presented in Figure 1.

[Table 1]

[Figure 1]

2.3. Quality and Risk of Bias Assessment

Two reviewers independently assessed the quality and risk of bias for each study using the Quality Assessment Tool for Observational Cohort and Cross-sectional Studies and the Cochrane Handbook for Systematic Reviews of Interventions for Controlled Trials (Higgins, Altman & Sterne, 2017). Quality ratings were highly consistent and did not require mediation by a third reviewer. See Appendix C for quality assessment summary tables.

2.4. Data Analysis

In addition to a narrative synthesis of all reviewed studies, a random effects meta-analysis was conducted using the statistical package Meta-Analysis via Shiny (MAVIS; Hamilton, 2011). Studies that examined the predictive value of early response and which reported sufficient statistical information to calculate standardised effect sizes were included in the meta-analysis. Hedges *g* was calculated to adjust for unequal sample size across groups, comparing post-treatment outcomes between early responders versus cases without early response. Heterogeneity was examined using the Q and I^2 statistics (Higgins et al., 2003). Potential publication bias was examined using a weight-function model (Vevea & Hedges, 1995) as well as the fail-safe N calculation using Rosenthal's method (Rosenthal, 1991).

There were several sources of heterogeneity in the design of the studies, including in the symptom measures used, the time-point at which early response was defined, and the time-point at which outcomes were assessed (e.g. post-treatment, at 6-months follow-up). We therefore followed a principled and consistent approach to enable meta-analysis, where we selected the outcome measure with the largest effect size (in studies with multiple outcomes), at the measurement point closest to session 4 (in studies with multiple measurements), with the outcome defined at the end of treatment (or most proximal follow-up assessment). In a subgroup analysis, we separately meta-analysed samples using measures for depression or anxiety symptoms. Other sources of heterogeneity such as differences in the types of treatments and study design were examined by removing systematically different studies from the meta-analysis and examining the influence on pooled effect sizes and indices of heterogeneity. The small number of studies (<20) that were included in meta-analysis and the imbalance in the number of studies per each of the above categories of heterogeneity precluded the use of categorical moderator analysis, according to conventional guidelines (Rubio-Aparicio, Sánchez-Meca, López-López, Botella, & Marín-Martínez, 2017).

[Table 2]

3. RESULTS

3.1. Study characteristics

Design and sample characteristics. Twenty-five studies met the inclusion criteria and are described in Tables 2 and 3, including analyses of controlled trials (*n*= 16) and observational studies (*n*= 9). Many studies focused exclusively on depression (*n*= 16), whilst only three focused exclusively on anxiety disorders. Sample sizes ranged from 23 to 5484 participants, with nearly all of the studies including more females than males.

Measures. A wide range of measures were used to assess early response, as shown in Table 2. The Beck Depression Inventory (BDI) and the Hamilton rating scale for depression (HAMD) were most commonly used to assess depression symptoms. A variety of different measures were used to assess anxiety, including the Beck Anxiety Inventory (BAI), Penn State Worry Questionnaire, and Panic Disorder Severity Scale-self report. Measures of psychological distress included the Outcomes Questionnaire (OQ-45), and the Clinical Outcomes in Routine Evaluation (CORE). Only two studies used general psychological distress measures rather than disorder-specific measures to assess symptom change (Lutz et al., 2017; Rubel et al., 2015), whilst two used a combination of both (Lutz, Stulz & Köck, 2009; Schibbye et al., 2014).

Defining early response. Many studies conceptualised early response by quantifying changes between intake assessments and week four of treatment, as reported by 15 of the 25 studies, while others conceptualised this as far as the 8th week. Sixteen studies used standard

and similar criteria to define early response, including reliable change (RC; Jacobson & Traux, 1991) a specific percentage reduction on the measures used (e.g., >25% reduction of symptom severity from baseline to the week of early response), or early remission of symptoms based on established diagnostic cut-off scores. Other studies defined early response using methods such as hierarchical agglomerative cluster analysis (HAC) or Growth Mixture Modelling (GMM) to identify different trajectories of change among participants. The three studies measuring sudden gains used Tang and DeRubeis' (1999b) definition, and each study modified the criteria to fit the measures used and the diagnosis of patients. With regards to the relationship between early response and outcomes, 18 of the 25 studies clearly stated the methods used to predict outcomes among early responders. These methods included receiver operating characteristic curves (ROC), logistic regression, general linear models (GLM), GMM, trajectory analysis, linear modelling of growth curves, cluster analysis, piecewise growth mixture modelling (PGMM) and HAC.

Interventions. Most of the studies focused solely on psychological interventions (*n*= 15), where cognitive behavioural therapy (CBT) was the most commonly studied intervention. Ten studies included samples that accessed either pharmacotherapy or psychological therapy (both treatments were combined in 3 of these studies).

Of the studies using exclusively psychological therapies, two used internet-based cognitive-behavioural therapy, one used group telephone therapy, and another pooled data for patients who accessed individual, group and internet-delivered therapies. Other psychological therapies included cognitive behavioural analysis system of psychotherapy (CBASP), supportive-expressive dynamic psychotherapy, interpersonal psychotherapy (IPT), behavioural activation, metacognitive therapy, short-term psychodynamic supportive psychotherapy (SPSP), behavioural therapy and self-control therapy. Thirteen studies used a variety of psychological or pharmacological treatments but not all of them provided

information about each type of therapy, and the different therapies were not all analysed separately.

3.2. Quality Assessment

Overall, most reviewed studies had low risk of bias. More than half (15 = 60%) were rated as "good" quality studies and a further 8 (32%) were rated as "fair". Only two were rated as "poor" quality studies with high risk of bias. The most common sources of bias included the failure to report dropouts / excluded cases, lack of recruitment information, little or no description of treatments, and lack of information about randomization procedures and blinding in clinical trials. A detailed risk of bias assessment for each study can be found in supplementary Appendix C.

3.3. Early response and treatment outcomes

Pooling available data across studies indicated that approximately half (47.9%) of study participants were classified as early responders. The majority of these studies found support for a positive association between early response and post-treatment outcomes as summarised in Table 3. Some studies additionally investigated the influence of pre-treatment changes that may have occurred prior to exposure to therapy (e.g., between intake assessments and the first therapy session). These studies found that early response to therapy remained a significant predictor after controlling for pre-treatment changes that might denote regression to the mean (Delgadillo et al., 2014; Lutz et al., 2017). Furthermore, early response was associated with clinical outcomes across different severities of psychological distress (e.g., mild, moderate, severe depression), different forms of psychotherapy / pharmacotherapy and even in a placebo control group (see: Tadić et al., 2010).

Only one study did not support an association between early response and outcomes, although they specifically focused on dropouts (Arnow et al., 2007). Two studies found

equivocal relationships between early response and positive outcomes. Başoğlu et al. (1994) found no association between early remission in panic symptoms and improvement in anticipatory anxiety and avoidance at post-treatment. However, they did find an association between early improvement in phobia symptoms and a reduction in panic attacks. Van et al. (2008) compared psychological therapy with a combination of pharmacotherapy and psychotherapy. In the psychotherapy condition, early non-response significantly predicted final non-response. However, this was not the case for remission rates. In the combined therapy condition, both final non-response and remission rates differed significantly for those who showed early non-response and early response. This may suggest that the medication in the combined therapy influenced the predictive value of early response. Therefore, the potential influence of combined treatments was examined using a subgroup meta-analysis described in the next section.

Some of the reviewed studies did not specifically test early response – outcome associations, but instead focused on when the largest changes occurred, how many cases showed early response, or which variables predicted early response (Gildengers et al., 2005; Heckman et al., 2017; Jordan et al., 2014; Rabin, Kaslow & Rehm, 1984). The key findings from these studies are summarised in Table 3.

[Figure 2]

3.4. Meta-analysis

A primary meta-analysis was conducted using effect size data from 15 studies (including 18 unique samples and 3956 participants) to examine associations between early response and post-treatment outcomes. The weighted mean effect size was $g = 0.87$ [95 % CI: 0.63, 1.10] p

< .0001, indicating that early responders had significantly better treatment outcomes compared to cases without early response (see Figure 2). Cochrane's Q-test revealed evidence of significant heterogeneity in the sample ($Q(17) = 246.12$, $p < .0001$) and the I^2 statistic was 87.73%, indicating a large degree of heterogeneity. The weight-function model likelihood ratio test did not indicate evidence of publication bias; $x^2(1) = 3.08$, $p = .08$. Calculation of the fail-safe N indicated that 3200 studies with non-significant results would be needed to conclude that the effect of early response is not statistically significant.

A series of subgroup analyses were conducted to examine the potential influence of methodological features. Removing samples ($k = 6$) with combined (psychotherapy and pharmacological) treatments, samples ($k = 4$) from studies that applied latent clustering of symptom trajectories, samples ($k = 4$) from studies that investigated sudden gains, and samples ($k = 6$) from studies with "fair" (instead of "good") quality ratings made little difference to the pooled effect sizes ($g = 0.82$ to 0.92) and indices of heterogeneity ($I^2 =$ 87.46 to 91.83%) obtained in subsamples. However, removing samples ($k = 4$) from observational studies considerably reduced the pooled effect size ($g = .74$) and indices of heterogeneity ($I^2 = 79.69\%$; $Q(13) = 66.72$, $p < .0001$) observed in the subgroup of RCT samples. Removing two studies (Başoğlu et al., 1994; Steinman et al., 2013) that specifically examined panic disorder interventions also considerably reduced the pooled effect size ($g =$ .77) and indices of heterogeneity ($I^2 = 79.70\%$; $Q(15) = 230.87$, $p < .0001$) observed in the subgroup of RCT samples. Furthermore, separate meta-analyses in the samples analysing depression ($k = 15$, n = 3690) and anxiety ($k = 5$, n = 2214) measures revealed markedly different effect sizes (depression $g = 0.76$ [0.58, 0.94]; anxiety $g = 1.37$ [0.86, 1.88]) and indices of heterogeneity (depression $I^2 = 76.27\%$; anxiety $I^2 = 90.00\%$). Forest plots for these separate meta-analyses are shown in supplementary Appendix D.

4. DISCUSSION

4.1. Summary of evidence

This systematic review provides robust evidence of a positive association between early response to psychological therapy and post-treatment outcomes. A large and statistically significant overall effect size (g = 0.87) indicated that early responders tended to have better treatment outcomes compared to other patients. Translating this effect size into an odds ratio scale (4.84), would indicate that early responders were at least four times more likely to attain positive treatment outcomes compared to other cases. This association was replicated across various forms of therapy including CBT, psychodynamic and interpersonal psychotherapies, and guided self-help interventions delivered in person, in groups, via telephone and via internet. Furthermore, there was low risk of bias in most of the reviewed studies, which were generally of good methodological quality, and we found no evidence of publication bias.

The effect of early response on anxiety outcomes was considerably larger (g = 1.37) compared to the effect on depression outcomes (g = 0.76). However, the index of heterogeneity for anxiety studies was much higher (I^2 = 90% vs. 76%), suggesting that the influence of early response may vary across anxiety disorders. In particular, large effect sizes (g = 1.80 to 1.87) were observed in the studies that examined early response to panic disorder treatments (Başoğlu et al., 1994; Steinman et al., 2013), and the exclusion of these two studies reduced the overall heterogeneity of meta-analytic findings. The design of studies was another relevant source of heterogeneity, as we found that the index of heterogeneity was somewhat attenuated in the subsample of randomised controlled trials (after excluding observational studies). This is plausibly explained by the stringent inclusion/exclusion criteria applied in clinical trials, which typically yields more diagnostically homogeneous samples compared to naturalistic observational studies.

4.2. Methodological Considerations

The studies included in this review varied considerably in the methods used to assess early response. For example, Rubel et al. (2015) compared three different approaches: growth mixture modelling (GMM), reliable change (RC) and clinical improvement. They found that at session three approximately 35% of the participants met criteria for RC, and 16.3% had achieved clinically significant improvement. On the other hand, when using GMM, only 7.2% showed early positive response by session three. This shows that GMM is a more conservative approach, identifying much fewer patients as showing early response than the RC and clinical improvement methods. This was also found when comparing the studies by Gilboa-Shechtman and Shahar (2006) and Lutz et al. (2009) which re-analysed the same dataset (Elkin et al., 1989). Although both studies included the same number of participants from the same dataset, they used different methods to identify early response and to predict outcomes. Consistent with Rubel et al. (2015), Lutz et al. (2009) identified fewer patients as showing early response (n=99) when using GMM, compared to Gilboa-Shechtman and Shahar (2006) who used linear modelling of growth curves (n=122).

Early response was most often defined by week 4, although other cut-offs were used in different studies. For example, Steidtmann et al. (2013) used weeks four, six and eight as they are considered to be clinically useful decision points. Similarly, Schibbye et al. (2014) used week five as it represented the first half of the treatment. The studies by Arnow et al. (2007), Başoğlu et al (1994) and Tadić et al. (2010) defined early response by week two of treatment. Alternatively, some studies retrospectively identified a week of early response via the use of statistical methods that modelled latent trajectories of change (e.g., Heckman et al., 2017; Lutz et al., 2009; Steinman et al., 2012). In spite of these methodological differences, subgroup analyses removing latent trajectory profiling studies from meta-analysis did not considerably change the pooled effect size or indices of heterogeneity. However, Rubel et al.

(2015) argue that a minimum of three sessions are required to conduct GMM, suggesting that there may not have been enough information within the first two weeks of treatment to assess the predictive value of early response. Therefore, it may be that week two is too early to find significant differences between those who do and do not show early response, which may possibly explain the mixed results found by Arnow et al. (2007), Başoğlu et al (1994) and Tadić et al. (2010). Overall, studies that defined early response by week four or later consistently found associations with post-treatment outcomes.

In total, only 10 of the 25 (40%) studies clearly stated how they handled missing data in their analyses. Not appropriately managing dropouts can lead to biased interpretations of results (Dumville, Torgerson & Hewitt, 2006; Mallinckrodt et al., 2001). Therefore, it may be that including and examining all participants who dropped out of treatment could potentially influence the results concerning early response and treatment outcomes. However, reviewed studies that used an intention-to-treat approach still found large and statistically significant effects of early response over treatment outcomes.

4.3. Strengths and limitations

To our knowledge, this is the first meta-analytic review to assess the association between early response and psychological therapy outcomes. Strengths of this review include the prospective registration of the review protocol and the study selection across multiple databases, which had no date restriction in order to gather all eligible studies. Reverse and forward citation searches were also conducted. The review also included a quality assessment by two independent assessors. However, the review excluded grey literature, studies published in languages other than English, and authors of eligible studies were not contacted to request data.

Some limitations should be taken into account when interpreting the findings of this review. Due to such a wide variety of measures and methods used to assess early response, directly comparing findings was difficult. Although a series of subgroup analyses were carried out, there was still a large amount of heterogeneity left unexplained. This is likely to have been due to the large number of measures used across the studies, with only two studies in the meta-analysis using the same depression measure. This meant that a subgroup analysis to assess the effect of specific measures was not possible. We also note that only 60% of reviewed studies provided sufficient information to compare treatment outcomes between cases with and without early response. In spite of the limited number of studies that contributed to meta-analysis, there was no evidence of publication bias and the fail-safe N (3200) indicated that the meta-analytic results were robust.

Sudden gains that occur early in treatment arguably overlap with the early response phenomenon. On this basis, we imposed inclusion criteria to ensure that we only analysed data from studies where the median sudden gain session occurred by week 4. This *a priori* exclusion criterion resulted in a stricter exclusion of sudden gains studies and a more liberal inclusion of early response studies. We argue that this is appropriate, given the focus of the current review. A more liberal inclusion of sudden gains studies would risk confounding these constructs, making it difficult to ascertain whether or not early response predicts better treatment outcomes (over and above the already known effect of sudden gains). We note that a subgroup analysis excluding sudden gains studies made no difference to the pooled meta-analysis results, so the inclusion of these studies is unlikely to bias results in any way.

Whilst the inclusion of RCTs and observational studies allowed for the inclusion of a greater number of studies, there are limitations in both designs. Observational studies have the advantage of including patients who are encountered in routine practice, thus strengthening the external validity of results. However, internal validity can be compromised

in observational studies, since they often lack rigorous features that are present in RCTs, such as the use of control groups, randomisation and blinding, adherence to treatment manuals, and stringent inclusion and exclusion criteria. RCTs included in this review also varied considerably in their patient selection criteria and how well they handled missing data. Although the study design accounted for a proportion of heterogeneity (as described above), it is nevertheless clear that early response is a phenomenon that occurs across a wide variety of study designs and settings. Another point to consider is that a large number of secondary analyses were included in the review. These studies should be interpreted with caution as secondary analyses can lack statistical power and multiple testing can inflate the chances of type 1 error.

4.4. Implications for research, theory and practice

It is important to consider the findings of this review in the context of previous research. Although sudden gains are conceptually different from early response, unless they occur by week four of therapy, they are both important for predicting post-treatment outcomes. A meta-analysis conducted by Aderka, Nickerson, Bøe and Hofmann (2012) found results consistent with the current review. They found robust evidence for a positive association between sudden gains and primary outcomes in depression and anxiety symptoms. However, whilst the current review focused on primary symptoms of depression and anxiety, Aderka et al. (2012) also examined secondary symptoms. Compared to the large effect size in primary symptoms (g=0.62, [0.43, 0.80]), the effect size for secondary symptoms was small and non-significant (g=0.37, [-0.34, 1.07]). This suggests that sudden gains in primary symptoms may not necessarily lead to improvements in secondary symptoms. As the current review did not address this, future research or reviews into early response should also focus on secondary symptoms. Secondly, Aderka et al. (2012) found effect sizes of sudden gains to be much larger in CBT interventions compared to non-CBT interventions. Only eight of the 25 studies

in the current review used therapies other than some form of cognitive therapy. This highlights the need for more research into early response in treatments other than cognitive therapies. Additionally, Aderka et al. (2012) found no differences in effect sizes across depression and anxiety, while the present review indicates a stronger influence of early response on anxiety outcomes compared to depression.

Previous authors have suggested that the early response effect could be explained by a trend of improvement that began prior to the start of treatment, which might be indicative of regression to the mean (Beckham, 1989). Two of the reviewed studies did in fact observe that early pre-treatment gains are associated with trajectories of improvement, however both studies still found significant effects of early response during treatment after controlling for pre-treatment gains (Delgadillo et al., 2014; Lutz et al., 2017). On this basis, the effect of early response cannot be entirely explained by regression to the mean. Furthermore, the pooled rate of early response across studies (47.9%) was considerably higher than the rate of short-term remission observed in waitlist control groups (~20%) of depression treatment trials (Posternak & Miller, 2001; Whiteford et al., 2013); hence the early response effect cannot simply be explained by spontaneous remission.

An alternative explanation could be derived from the findings by Tadić et al. (2010), which indicated that individuals in a placebo control group also experienced early response at an equal or higher rate than the CBT group and with similar effects on the eventual remission of symptoms. This could potentially be explained by the nonspecific factors discussed by Ilardi and Craighead (1994), who suggest that early response may simply be due to a "readiness to change", and is therefore a quasi-placebo response to factors such as increased hope and expectations of improvement. Howard et al. (1993) have previously referred to this as a *remoralisation* effect which they attributed to common factors that are likely to be present in most forms of psychotherapy. This explanation fits within the wider literature on

the placebo effect in healthcare. For example, prior studies have shown experimentally that perceiving clinicians as warm, friendly, empathic and competent enhances the placebo effect (Howe, Goyer, & Crum, 2017; Kaptchuk et al., 2008). However, given that the four other studies in this review using a control condition either only reported early response in the treatment condition or pooled the data, it was not possible to determine whether early response occurs consistently across the placebo conditions. Future individual patient data meta-analyses of trials comparing psychotherapies with placebo could help to elucidate the extent to which the early response phenomenon may be partly or fully explained by a quasi-placebo effect.

An important clinical implication of this literature is that response to treatment during the first month is a reliable prognostic indicator across multiple forms of psychological care for common mental health problems. The fourth session could be seen as a timely opportunity to review treatment outcomes and to identify potential obstacles to improvement, which is consistent with the principles of routine outcome monitoring and feedback-informed treatment (Carlier et al., 2012; Lambert et al. 2003).

4.5. Conclusions

In conclusion, a robust evidence-base supports the notion that patients showing reliable symptomatic improvements within the initial therapy sessions tend to have a much better prognosis after treatment. Therefore, routinely monitoring early response is of paramount importance, as it has the potential to identify those who may have a protracted or limited response to treatment. A clinical review using validated outcome measures at the fourth session of therapy is recommended, in order to assess progress and to identify potential obstacles to improvement in cases that have not yet shown reliable symptomatic improvements.

DATA AVAILABILITY STATEMENT

All relevant data sources will be made available upon written request to the corresponding author.

CONFLICT OF INTERESTS

The authors declare no conflict of interests.

REFERENCES

Aderka, I. M., Nickerson, A., Bøe, H. J., & Hofmann, S. G. (2012). Sudden gains during psychological treatments of anxiety and depression: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *80*(1), 93-101.

Arnow, B. A., Blasey, C., Manber, R., Constantino, M. J., Markowitz, J. C., Klein, D. N., ... & Rush, A. J. (2007). Dropouts versus completers among chronically depressed outpatients. *Journal of Affective Disorders*, *97*(1), 197-202.

Başoğlu, M., Marks, I. M., Kiliç, C., Swinson, R. P., Noshirvani, H., Kuch, K., & O'Sullivan, G. (1994). Relationship of panic, anticipatory anxiety, agoraphobia and global improvement in panic disorder with agoraphobia treated with alprazolam and exposure. *The British Journal of Psychiatry*, *164*(5), 647-652.

Beckham, E. E. (1989). Improvement after evaluation in psychotherapy of depression: Evidence of a placebo effect? *Journal of Clinical Psychology*, *45*(6), 945-950.

Bell, C., Waller, G., Shafran, R., & Delgadillo, J. (2017). Is there an optimal length of psychological treatment for eating disorder pathology? *International Journal of Eating Disorders, 50*(6), 687-692.

Bradford, A., Cully, J., Rhoades, H., Kunik, M., Kraus-Schuman, C., Wilson, N., & Stanley, M. (2011). Early response to psychotherapy and long-term change in worry symptoms in older adults with generalized anxiety disorder. *The American Journal of Geriatric Psychiatry*, *19*(4), 347-356.

Braun, J. D., Strunk, D. R., Sasso, K. E., & Cooper, A. A. (2015). Therapist use of Socratic questioning predicts session-to-session symptom change in cognitive therapy for depression. *Behaviour Research and Therapy*, *70*, 32-37.

Carlier, I. V., Meuldijk, D., Van Vliet, I. M., Van Fenema, E., Van der Wee, N. J., & Zitman, F. G. (2012). Routine outcome monitoring and feedback on physical or mental health status: Evidence and theory. *Journal of Evaluation in Clinical Practice, 18*, 104 –110.

Crits-Christoph, P., Connolly, M. B., Gallop, R., Barber, J. P., Tu, X., Gladis, M., & Siqueland, L. (2001). Early improvement during manual-guided cognitive and dynamic psychotherapies predicts 16-week remission status. *The Journal of Psychotherapy Practice and Research*, *10*(3), 145-154.

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.

Dew, M. A., Reynolds, C. F., Houck, P. R., Hall, M., Buysse, D. J., Frank, E., & Kupfer, D. J. (1997). Temporal profiles of the course of depression during treatment: predictors of pathways toward recovery in the elderly. *Archives of General Psychiatry*, *54*(11), 1016-1024.

Doyle, P. M., Le Grange, D., Loeb, K., Doyle, A. C., & Crosby, R. D. (2010). Early response to family-based treatment for adolescent anorexia nervosa. *International Journal of Eating Disorders*, *43*(7), 659-662.

Dumville, J. C., Torgerson, D. J., & Hewitt, C. E. (2006). Research methods: reporting attrition in randomised controlled trials. *BMJ: British Medical Journal*, *332*(7547), 969-971.

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., ... & Fiester, S. J. (1989). National Institute of Mental Health treatment of depression collaborative research program: General effectiveness of treatments. *Archives of General Psychiatry*, *46*(11), 971-982.

Gilboa-Schechtman, E., & Shahar, G. (2006). The sooner, the better: Temporal patterns in brief treatment of depression and their role in long-term outcome. *Psychotherapy Research*, *16*(3), 374-384.

Gildengers, A. G., Houck, P. R., Mulsant, B. H., Dew, M. A., Aizenstein, H. J., Jones, B. L., ... & Reynolds III, C. F. (2005). Trajectories of treatment response in late-life depression: psychosocial and clinical correlates. *Journal of Clinical Psychopharmacology*, *25*(4), S8-S13.

Gois, C., Dias, V. V., Carmo, I., Duarte, R., Ferro, A., Santos, A. L., ... & Barbosa, A. (2014). Treatment response in type 2 diabetes patients with major depression. *Clinical Psychology & Psychotherapy*, *21*(1), 39-48.

Grilo, C. M., Masheb, R. M., & Wilson, G. T. (2006). Rapid response to treatment for binge eating disorder. *Journal of Consulting and Clinical Psychology*, *74*(3), 602-613.

Haas, E., Hill, R. D., Lambert, M. J., & Morrell, B. (2002). Do early responders to psychotherapy maintain treatment gains? *Journal of Clinical Psychology*, *58*(9), 1157-1172.

Hamilton, W. (2011). Package 'MAVIS' (1st ed.). Retrieved from https:// www.researchgate.net/publication/280711501_MAVIS_Meta_Analysis _via_Shiny

Heckman, T. G., Heckman, B. D., Anderson, T., Bianco, J. A., Sutton, M., & Lovejoy, T. I. (2017). Common factors and depressive symptom relief trajectories in group teletherapy for persons ageing with HIV. *Clinical Psychology & Psychotherapy*, *24*(1), 139-148.

Higgins, J. T., Altman, D. G., & Sterne, J. A. C. (2011). Assessing risk of bias in included studies. In J. T. Higgins & S. Green (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions version 5.1.0 (updated March 2011; Chap.8).* London, England: The Cochrane Collaboration. Retrieved from https://www.handbook.cochrane.org

Howard, K. I., Lueger, R. J., Maling, M. S., and Martinovich, Z. (1993). A phase model of psychotherapy outcome: causal mediation of change. *Journal of Consulting and Clinical Psychology, 61*(4), 678-685.

Howe, L. C., Goyer, J. P., & Crum, A. J. (2017). Harnessing the placebo effect: Exploring the influence of physician characteristics on placebo response. *Health Psychology, 36*(11), 1074.

Hunnicutt-Ferguson, K., Hoxha, D., & Gollan, J. (2012). Exploring sudden gains in behavioral activation therapy for major depressive disorder. *Behaviour Research and Therapy*, *50*(3), 223-230.

Ilardi, S. S., & Craighead, W. E. (1994). The role of nonspecific factors in cognitive-behavior therapy for depression. *Clinical Psychology: Science and Practice*, *1*(2), 138-155.

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.

Jordan, J., Carter, J. D., McIntosh, V. V., Fernando, K., Frampton, C. M., Porter, R. J., ... & Joyce, P. R. (2014). Metacognitive therapy versus cognitive behavioural therapy for depression: a randomized pilot study. *Australian & New Zealand Journal of Psychiatry*, *48*(10), 932-943.

Kaptchuk, T.J., Kelley, J.M., Conboy, L.A., Davis, R.B., Kerr, C.E., Jacobson, E.E., Kirsch, I., Schyner, R.N., Nam, B.H., Nguyen, L.T. and Park, M. (2008). Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. *BMJ, 336*(7651), 999-1003.

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice, 10*, 288 –301.

Lewis, C. C., Simons, A. D., & Kim, H. K. (2012). The role of early symptom trajectories and pretreatment variables in predicting treatment response to cognitive behavioral therapy. *Journal of Consulting and Clinical Psychology*, *80*(4), 525-534.

Lutz, W., Arndt, A., Rubel, J., Berger, T., Schröder, J., Späth, C., ... & Fuhr, K. (2017). Defining and Predicting Patterns of Early Response in a Web-Based Intervention for Depression. *Journal of Medical Internet Research*, *19*(6), e206.

Lutz, W., Stulz, N., & Köck, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *Journal of Affective Disorders*, *118*(1), 60-68.

Mallinckrodt, C. H., Clark, W. S., & David, S. R. (2001). Accounting for dropout bias using mixed-effects models. *Journal of Biopharmaceutical Statistics*, *11*(1-2), 9-21.

Masterson, C., Ekers, D., Gilbody, S., Richards, D., Toner-Clewes, B., & McMillan, D. (2014). Sudden gains in behavioural activation for depression. *Behaviour Research and Therapy*, *60*, 34-38.

Posternak, M. A., & Miller, I. (2001). Untreated short-term course of major depression: a meta-analysis of outcomes from studies using wait-list control groups. *Journal of Affective Disorders*, *66*(2), 139-146.

Rabin, A. S., Kaslow, N. J., & Rehm, L. P. (1984). Changes in symptoms of depression during the course of therapy. *Cognitive Therapy and Research*, *8*(5), 479-487.

Rosenthal, R. (1991). *Meta-analytic Procedures for Social Research* (Vol. 6). Sage.

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), 478-488.

Rubio-Aparicio, M., Sánchez-Meca, J., López-López, J. A., Botella, J., & Marín-Martínez, F. (2017). Analysis of categorical moderators in mixed-effects meta-analysis: Consequences of using pooled versus separate estimates of the residual between-studies variances. *British Journal of Mathematical and Statistical Psychology, 70*(3), 439-456.

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.

Schibbye, P., Ghaderi, A., Ljótsson, B., Hedman, E., Lindefors, N., Rück, C., & Kaldo, V. (2014). Using early change to predict outcome in cognitive behaviour therapy: exploring timeframe, calculation method, and differences of disorder-specific versus general measures. *PloS one*, *9*(6), e100614.

Stauffer, V. L., Case, M., Kinon, B. J., Conley, R., Ascher-Svanum, H., Kollack-Walker, S., ... & Lieberman, J. (2011). Early response to antipsychotic therapy as a clinical marker of subsequent response in the treatment of patients with first-episode psychosis. *Psychiatry Research*, *187*(1), 42-48.

Steidtmann, D., Manber, R., Blasey, C., Markowitz, J. C., Klein, D. N., Rothbaum, B. O., ... & Arnow, B. A. (2013). Detecting critical decision points in psychotherapy and psychotherapy+ medication for chronic depression. *Journal of Consulting and Clinical Psychology*, *81*(5), 783-792.

Steinman, S. A., Hunter, M. D., & Teachman, B. A. (2013). Do patterns of change during treatment for panic disorder predict future panic symptoms? *Journal of Behavior Therapy and Experimental Psychiatry*, *44*(2), 150-157.

Tadić, A., Helmreich, I., Mergl, R., Hautzinger, M., Kohnen, R., Henkel, V., & Hegerl, U. (2010). Early improvement is a predictor of treatment outcome in patients with mild major, minor or subsyndromal depression. *Journal of Affective Disorders*, *120*(1), 86-93.

Tang, T. Z., & DeRubeis, R. J. (1999a). Reconsidering rapid early response in cognitive behavioral therapy for depression. *Clinical Psychology: Science and Practice*, *6*(3), 283-288.

Tang, T. Z., & DeRubeis, R. J. (1999b). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology, 67*(6), 894-904.

Van, H. L., Schoevers, R. A., Kool, S., Hendriksen, M., Peen, J., & Dekker, J. (2008). Does early response predict outcome in psychotherapy and combined therapy for major depression? *Journal of Affective Disorders*, *105*(1), 261-265.

van Calker, D., Zobel, I., Dykierek, P., Deimel, C. M., Kech, S., Lieb, K., ... & Schramm, E. (2009). Time course of response to antidepressants: predictive value of early improvement and effect of additional psychotherapy. *Journal of Affective Disorders*, *114*(1), 243-253.

Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika, 60*(3), 419-435.

Whiteford, H. A., Harris, M. G., McKeon, G., Baxter, A., Pennell, C., Barendregt, J. J., & Wang, J. (2013). Estimating remission from untreated major depression: a systematic review and meta-analysis. *Psychological Medicine, 43*(8), 1569-1585.

Table 1. Inclusion and exclusion criteria

	Review question	
	Does early response to psychological therapy predict post-treatment depression or anxiety outcomes?	
	Inclusion criteria	**Exclusion criteria**
Population	Adult patients 18 years and over accessing psychological interventions for anxiety (including Obsessive-Compulsive Disorder, Post-Traumatic Stress Disorder, other anxiety disorders) and/or depression.	Studies exclusively with children and/or adolescents under 18. Studies exclusively with non-clinical samples.
Intervention	Any form of psychological therapy or intervention with the aim of treating depression or anxiety and delivered in any modality (individual, group, internet-based etc.).	Studies that do not include psychological interventions for depression and/or anxiety.
Comparator	A between-groups comparison of those who do and do not show early response to psychological therapy. Studies comparing treatment outcomes between cases that had sudden gains early in treatment (e.g. sample median sudden gain occurred within the first 4 weeks) and cases without sudden gains.	Studies where no comparisons are made between cases with and without early response to psychological treatment. Studies where sudden gains tended to occur after the fourth week of therapy.
Outcomes	The statistical significance and magnitude of the association between early response and post-treatment depression and anxiety outcomes.	Studies where early response is not measured. Studies where post-treatment depression / anxiety outcomes are not reported.
Setting	Any settings where psychological interventions are usually delivered, including online or telephone therapy, in any country.	
Study design	Randomised controlled trials and observational studies. Studies published in peer reviewed scientific journals, in the English language.	Grey literature, such as dissertation theses, which are not published in peer reviewed scientific journals. Editorials, newspaper or magazine articles and other forms of media. Literature sources not in published in English.

Table 2. Characteristics of studies included in the review

First Author and Year	Study Design	Study Setting	Primary Disorders	Analysed N	Intervention Condition	Outcome Measures[j]	Intervention Duration
Arnow et al. (2007)	RCT[b]	NA	MDD	681	CBASP, Nefazadone or COMB[h]	HAMD-24	12 weeks
Başoğlu et al. (1994)	RCT[b]	UK and US, 2 outpatient sites	PD, agoraphobia	154	AE, PE, AR or PR	Main phobia targets: avoidance and fear[k]	8 to 16 weeks[l]
Beckham. (1989)	Observational	US, university-based research site	Depression	23	CBT	BDI	NA
Bradford et al. (2011)	RCT[b]	US, Primary care clinic	GAD	76	CBT	PSWQ	12 weeks
Braun et al. (2015)	Observational[c]	NA	Depression	55	CT	BDI-II, HRSD	16 weeks
Crits-Christoph et al. (2001)	RCT[b]	NA	Mixed[e]	98[g]	CT, SE	BDI, BAI	16 to 52 weeks[m]
Delgadillo et al. (2014)	Observational[c]	UK, 1 primary care mental health service	Mixed[e]	1850	LiCBT	PHQ-9, GAD-7	Up to 10 weeks
Dew et al. (1997)[a]	Observational[c]	US, psychiatric institute and clinic	Depression	95	Combined NT and IPT	HAMD-17	18 weeks[n]
Gilboa-Schechtman et al. (2006)	RCT[b]	US, 3 university research sites	MDD	162	CBT, IPT, IMI-CM or PLA-CM	BDI, HRSD	16 weeks
Gildengers et al. (2005)[a]	RCT[b]	US, university based research site	MDD	395	IPT, NT or PX	HRSD-17	12 weeks
Gois et al. (2014)	RCT	Portugal, 3 outpatient diabetes clinics	MDD	30	IPT or Sertraline	MADRS	24 weeks
Heckman et al. (2017)	RCT[b]	US, telephone based	Depression	103	Group Teletherapy	GDS	12 weeks
Hunnicutt-Ferguson et al. (2012)[a]	Observational	NA	MDD[f]	42	BA	QIDS-SR	16 weeks
Jordan et al. (2014)	RCT pilot study	NZ, university clinical research unit	MDD, Bipolar II	48	MCT vs CBT	QIDS 16-C	12 weeks

Lewis et al. (2012)	Observational	US, university training clinic	Mixed[e]	173	CBT	BDI	NA
Lutz et al. (2009)	RCT[b]	US, 3 university research sites	MDD	162	CBT, IPT, IMI-CM or PLA-CM	BDI, HSCL-90	16 weeks
Lutz et al. (2017)	RCT	Internet-based[d]	Depression	409	ICBT	PHQ-9	12 weeks
Masterson et al. (2014)	RCT[b]	UK, mental health practices	Depression	40	BA	PHQ-9	12 weeks
Rabin et al. (1984)	RCT	NA	Depression	98	BT, CT or CBT and Self-Control Group Therapy	BDI	10 weeks
Rubel et al. (2015)	Observational	Germany, 26 counselling, medical or mental health sites	Mixed[e]	5484	Psychotherapy[i]	GMH	Not fixed[o]
Schibbye et al. (2014)	Observational	Sweden, intenet-based	Mixed[e]	112	ICBT	OQ-45, CORE-10, MADRS-S, PDSS-SR, LSAS-SR	10 weeks or 15 weeks[p]
Steidtmann et al. (2013)	RCT[b]	Outpatient sites[d]	MDD	352	CBASP or COMB[h]	IDS-SR	12 weeks
Steinman et al. (2013)	Observational[c]	US, university research clinic	PD	36	CBGT	PDSS	12 weeks
Tadić et al. (2010)	RCT[b]	Germany, outpatient facility	Depression	223	Sertraline or CBT	HAMD-17	10 weeks
Van et al. (2008)	RCT[b]	Netherlands, 2 outpatient facilities	Depression	190	SPSP or SPSP with FX or VX	HAMD-17	16 weeks

RCT, randomised control trial; NA, not available; UK, United Kingdom; US, United States; NZ, New Zealand; MDD, major depressive disorder; PD, panic disorder; GAD, generalised anxiety disorder; CBASP, cognitive behavioural analysis system of psychotherapy; AE, alprazolam plus exposure; PE, placebo plus exposure; AR, alprazolam plus relaxation; PR, placebo plus relaxation; CBT, cognitive behavioural therapy; LiCBT, low intensity CBT (guided self-help); CT, cognitive therapy; SE, supportive-expressive dynamic psychotherapy; NT, Nortriptyline; IPT, interpersonal psychotherapy; IMI-CM, imipramine plus clinical management; PLA-CM, placebo plus clinical management; PX, Paroxetine; BA, behavioural activation; MCT, meta-cognitive therapy; ICBT, internet cognitive behavioural therapy; BT, behavioural therapy; CBGT, cognitive behavioural group therapy; SPSP, short-term psychodynamic supportive psychotherapy; FX, Fluoxetine; VX, venlafaxine; HAMD, Hamilton rating scale for depression; BDI, Beck depression inventory; PSWQ, Penn state worry questionnaire; HRSD, Hamilton rating scale for depression; BAI, Beck anxiety inventory; MADRS, Montgomery-Åsberg depression rating scale; GDS, geriatric depression scale; QIDS-SR, quick inventory of depressive symptomatology-self report; QIDS-C, quick inventory of depressive symptomatology-Clinician assessed; HSCL, Hopkins symptom checklist; PHQ, patient health questionnaire; GMH, general mental health; OQ, outcomes questionnaire; CORE, clinical outcomes in routine evaluation; PDSS-SR, panic disorder severity scale-self report; LSAS-SR Liebowitz social anxiety scale-self report; IDS-SR; inventory of depressive symptomatology-self report; SG, sudden gain.

[a]Open trial studies, therefore the participants and therapists were not blind to treatment conditions.

[b]RCT refers to a secondary analysis of RCT.

[c]Observational refers to a secondary analysis of observational studies.

[d]Country of study not reported

[e]Mixed refers to a variety of depression and anxiety disorders.

[f]Healthy controls were also enrolled to take part in the study

[g]Here we report the largest sample reported by the authors, which was used to investigate the predictive value of early response.

[h]COMB refers to the psychotherapy used and the medication used combined
[i]Does not state what psychotherapy methods were used
[j]Only measures used to assess early response or early sudden gains
[k]Main phobia targets were modified from Marks and Mathews (1979). Includes avoidance, self and assessor rated, and fear, self and assessor rated.
[l]Psychological treatment lasted 8 weeks, medication could continue to week 16, 20 weeks for patients in the cross validation sample
[m]16 weeks for depression or anxiety with 3 monthly booster sessions, 52 weeks for cases with avoidant or obsessive-compulsive personality disorder
[n]The authors focused only on the 18 weeks of therapy before randomisation (secondary analysis of RCT)
[o]Treatment was not fixed to a strict time limit, sessions varied from 4 to 109 sessions (m=9.76, SD = 8.25)
[p]10 weeks for depression and panic disorder, 15 weeks for social anxiety disorder

Table 3. Findings reported by studies included in the review

First Author and Year	**Early Response Timing[f]**	**N (% of total)**		**Reported Statistics/Findings**
		Early Responders	**Others**	
Arnow et al. (2007)	Week 2	NA	NA	No difference between dropouts and completers in terms of early response (main effect F=.33, p=.57; dropout x treatment F=.11, p=.89).
Başoğlu et al, (1994)	Week 2-4	NA	NA	Early improvement (by week 4) in avoidance of main phobia targets explained 45% of variance in clinical global impression (CGI) ratings by week 8. Week-4 improvement in fear explained 24% of variance in patients' global impression (PGI) ratings by week 8.
Beckham. (1989)	Session 4-6	NA	NA	The BDI at session 6 correlated with the final session BDI (r=.80, p<.001).
Bradford et al. (2011)	Week 4	23 (30.3%)	53	Responders at 3 months had a significantly larger magnitude of change at 4 weeks compared to non-responders (t(74)=3.423, p<.001). This was the same for 15 month responders (t(74)=3.069, p=.003).
Braun et al. (2015)	Week 4	NA	NA	Symptom improvements from session 1 to 4 were significant and substantial (d=1.59, t=5.58, p<.001).
Crits-Christoph et al. (2001)	Week 4	NA	NA	Change on the BDI and BAI measures from baseline to weeks 2, 3 and 4 significantly predicted remission at week 16 (all p<.005). At week 4, the Area under the curve (AUC) values for early change in each outcome measure were .73 for BDI and .77 for BAI.
Delgadillo et al. (2014)	Week 4	NA	NA	For patients whose treatment lasted ≥5 sessions, reliable improvement by session 4 predicted post-treatment reliable and clinically significant improvement; with significant odds ratios in depression (pooled OR = 12.60, p < .001) and anxiety (pooled OR = 21.10, p < .001) measures.
Dew et al. (1997)	Week 4-6	29 (30.5%)	66	62% of the early response group already had below threshold HAMD scores by week 6. All subjects had these low scores by weeks 8-10. By week 4, early responders had a higher percentage of cases with remission of symptoms compared to non-early responders (x^2 = 13.1, p < .01).

Gilboa-Schechtman et al. (2006)	Week 4	122 (75.3%)	40	The end of treatment BDI (F = 18.17, p < .01) and HRSD (F = 22.59, p < .01) scores were significantly lower for the rapid responders group compared to others.
Gildengers et al. (2005)	Week 2-5	181 (47.9%)	197	Early responders in NT+IPT reached HRSD score of 10 by week 5, PX+IPT by week 3 and NT/PX by week 2.
Gois et al. (2014)	Week 6	22 (73.3%)	8	Early responders showed significant improvement in depression over time in both IPT and Sertraline groups (F(2.4)=27.73, p<.001).
Heckman et al. (2017)	Week 4	32 (30.5%)	73	Post-treatment scores for early responders M=7.43, SD=4.2, delayed responders M=6.41, SD=4.64, non-responders M=17.3, SD=6.17.
Hunnicutt-Ferguson et al. (2012)[b]	SG median pre-gain session 1[a]	15 (35.7%)	27	Sudden gain patients had significantly different rates of treatment response to non-sudden gain participants (X^2=5.09, p<.05).
Jordan et al. (2014)	Week 4	NA	NA	Participants in both therapies showed clinically significant improvements on the QIDS at week 4 (MCT d=.74, 95% CI [.30, 1.17]; CBT d=.73, 95% CI [.31, 1.14]).
Lewis et al. (2012)	Session 5	NA	NA	Around 76% of total symptom change was observed after intake, reaching 85.47% of total symptom change at session 5.
Lutz et al. (2009)	Week 8	99 (61.1%)	63	Class membership (early response and non-early response) was significantly associated with reliable improvement post-treatment (X^2(9)=74.8, p<.001).
Lutz et al. (2017)	Week4	343 (83.9%)	66	Rapid early response effect size d=2.9; moderate early improvement d=1.19.
Masterson et al. (2014)	SG median pre-gain session 2[a]	17 (42.5%)	23	At post-treatment sudden gain groups had a mean PHQ score of M=5.3, SD=3.63, whereas non-sudden gain participants had M=10.2, SD=6.8; t(35)=2.92, p=.006.
Rabin et al. (1984)	Week 3-4	NA	NA	For the majority of depressive symptoms, the largest changes occurred during the first 2 to 4 sessions.
Rubel et al. (2015)	Session 3	396 (7.2%)	5088	Early responders showed the highest pre-post effect sizes (ds=1.88-2.16) and the highest share of reliably improved patients at post-treatment (90-93%).

Schibbye et al. (2014)	Week 4	NA	NA	The disorder specific measures at week 4 explained variance in outcomes at post-treatment (panic disorder R^2=.34, p<.01; depression R^2=.41, p>.01; social anxiety disorder R^2=.43, p<.01).
Steidtmann et al. (2013)	Week 6-8	NA	NA	Percentage of symptom reduction at weeks 4, 6 and 8 predicted post-treatment HRSD remission status in combined treatment (week 4: x^2 = 6.26) and psychotherapy (week 4: x^2 = 4.05).
Steinman et al. (2013)	SG session 2	19 (52.8%)	17	Cluster membership (sudden gains versus others) was a significant predictor of post-treatment outcomes, accounting for 25% of the variability in six-month follow-up scores.
Tadić et al. (2010)	Week 2	95 (42.6%)	128	In both CBT and Sertraline, early response was a highly sensitive predictor of later stable response (76-82%) and stable remission (71-72%).
Van et al. (2008)	Week 8	109 (57.4%)	81	Early responders and non-responders in the psychotherapy condition were significantly different from each other in terms of final non-response (x^2=5.069, p<.01), but this was not the case for remission rates (x^2=.994, p=.319). In combined therapy, there were significant differences in terms of non-response (x^2=16.019, p<.001) and remission (x^2=12.435, p<.001).

SG, sudden gain; NA, not available. Overall mean proportion of cases with early response across studies = 47.9%

[a] Median pre-gain session refers to the therapy session immediately preceding the sudden gain (Tang & DeRubeis, 1999b)

Figure 1. PRISMA flow diagram of the systematic study selection

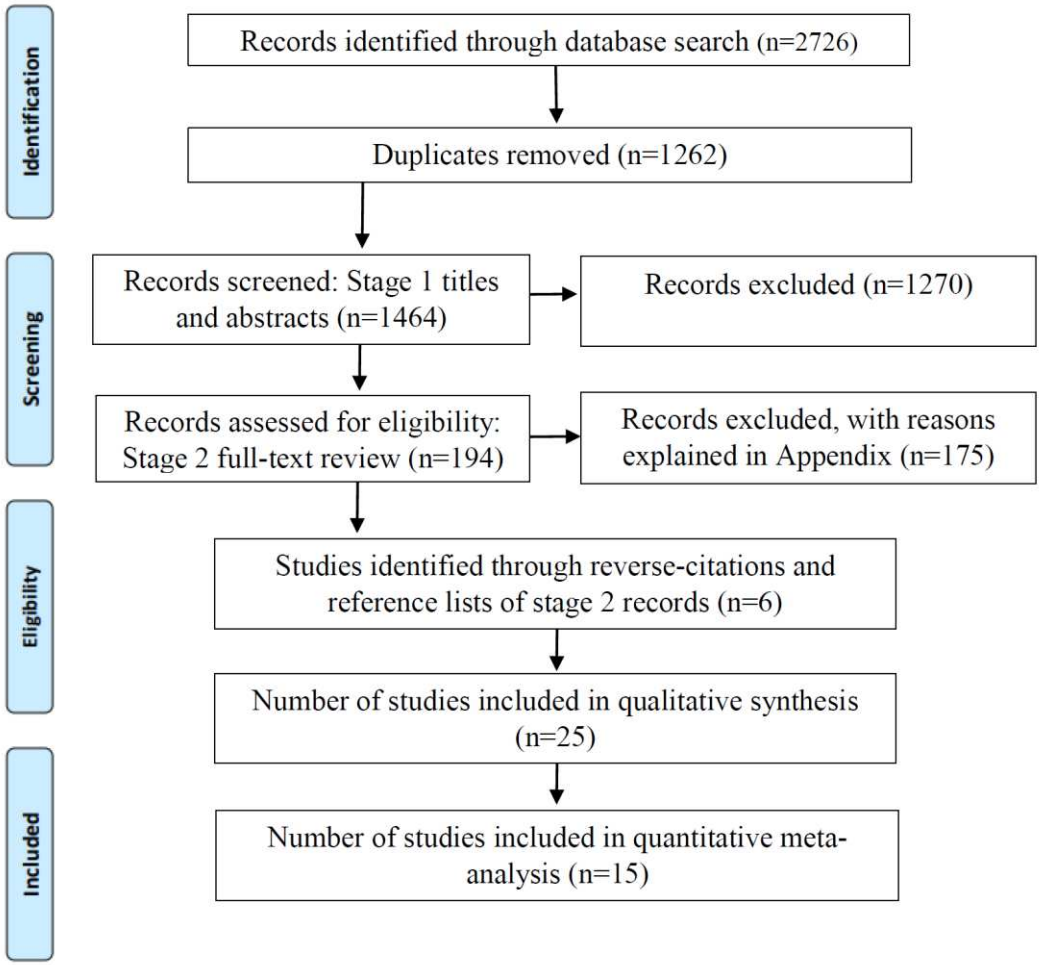

Figure 2. Random effects meta-analysis: post-treatment effect sizes comparing early responders vs. others