



In the Absence of Effects: An Individual Patient Data Meta-Analysis of Non-response and Its Predictors in Internet-Based Cognitive Behavior Therapy

Alexander Rozental^{1,2}, Gerhard Andersson^{1,3} and Per Carlbring^{4,5*}

¹ Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden, ² Institute of Child Health, University College London, London, United Kingdom, ³ Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden, ⁴ Department of Psychology, Stockholm University, Stockholm, Sweden, ⁵ Department of Psychology, University of Southern Denmark, Odense, Denmark

OPEN ACCESS

Edited by:

Osmano Oasi,
Catholic University of Sacred Heart,
Italy

Reviewed by:

Noelle Bassi Smith,
Veterans Health Administration (VHA),
United States

Sarah Bloch-Elkouby,
Mount Sinai Health System,
United States

Melissa Miléna De Smet,
Ghent University, Belgium

*Correspondence:

Per Carlbring
per.carlbring@psychology.su.se

Specialty section:

This article was submitted to
Clinical and Health Psychology,
a section of the journal
Frontiers in Psychology

Received: 05 November 2018

Accepted: 01 March 2019

Published: 29 March 2019

Citation:

Rozental A, Andersson G and
Carlbring P (2019) In the Absence of
Effects: An Individual Patient Data
Meta-Analysis of Non-response and
Its Predictors in Internet-Based
Cognitive Behavior Therapy.
Front. Psychol. 10:589.
doi: 10.3389/fpsyg.2019.00589

Background: Negative effects of psychological treatments have recently received increased attention in both research and clinical practice. Most investigations have focused on determining the occurrence and characteristics of deterioration and other adverse and unwanted events, such as interpersonal issues, indicating that patients quite frequently experience such incidents in treatment. However, non-response is also negative if it might have prolonged an ongoing condition and caused unnecessary suffering. Yet few attempts have been made to directly explore non-response in psychological treatment or its plausible causes. Internet-based cognitive behavior therapy (ICBT) has been found effective for a number of diagnoses but has not yet been systematically explored with regard to those patients who do not respond.

Methods: The current study collected and aggregated data from 2,866 patients in 29 clinical randomized trials of ICBT for three categories of diagnoses: anxiety disorders, depression, and other (erectile dysfunction, relationship problems, and gambling disorder). Raw scores from each patient variable were used in an individual patient data meta-analysis to determine the rate of non-response on the primary outcome measure for each clinical trial, while its potential predictors were examined using binomial logistic regression. The reliable change index (RCI) was used to classify patients as non-responders.

Results: Of the 2,118 patients receiving treatment, and when applying a RCI of $z \geq 1.96$, 567 (26.8%) were classified as non-responders. In terms of predictors, patients with higher symptom severity on the primary outcome measure at baseline, Odds Ratio (OR) = 2.04, having a primary anxiety disorder (OR = 5.75), and being of male gender (OR = 1.80), might have higher odds of not responding to treatment.

Conclusion: Non-response seems to occur among approximately a quarter of all patients in ICBT, with predictors related to greater symptoms, anxiety disorders, and

gender indicating increasing the odds of not responding. However, the results need to be replicated before establishing their clinical relevance, and the use of the RCI as a way of determining non-response needs to be validated by other means, such as by interviewing patients classified as non-responders.

Keywords: negative effects, non-response, predictors, individual patient data meta-analysis, Internet-based cognitive behavior therapy

INTRODUCTION

Negative effects of psychological treatments are a relatively uncharted territory in both research and clinical practice. Despite being recognized early in the scientific literature (c.f., Strupp and Hadley, 1977), empirical evidence for their occurrence and characteristics have been quite scarce, but has currently received increased attention (Rozenal et al., 2018). Bergin (1966) provided the first report of the “client-deterioration phenomenon” (p. 236), referred to as the deterioration effect, i.e., patients faring worse in treatment. Since then, several studies have investigated the rate of worsening in different naturalistic settings (c.f., Hansen et al., 2002; Mechler and Holmqvist, 2016; Delgado et al., 2018), while a number of systematic reviews have assessed deterioration among patients in randomized controlled trials (c.f., Ebert et al., 2016; Rozenal et al., 2017; Cuijpers et al., 2018), estimating that ~5–10% of those in treatment for depression and anxiety disorders deteriorate. In comparison to a wait-list control, the odds ratio for deterioration in treatment is nevertheless lower, suggesting that the benefits of receiving help still outweigh the risks (Karyotaki et al., 2018). Recent attempts to identify variables related to worsening have also revealed that sociodemographics variables like lower educational level are linked to increased odds of deterioration (Ebert et al., 2016), while older age and having a relationship constitute protective factors (Rozenal et al., 2017). This implies that certain features might be important to consider in relation to treatment, although more research is needed to determine if and how this could be clinically useful.

Meanwhile, others have stressed the importance of monitoring the potential adverse and unwanted events that may occur in treatment, which are not necessarily related to symptoms (Mays and Franks, 1980). This can include interpersonal issues, stigma, and feelings of failure, identified using therapist checklists (Linden, 2013), self-reports completed by patients (Rozenal et al., 2016), or open-ended questions (Rozenal et al., 2015). Such incidents have been even less explored, although a few recent attempts have found that almost half of the patients are experiencing negative effects at some time in treatment (Rheker et al., 2017; Moritz et al., 2018; Rozenal et al., in press). Whether or not these are in fact detrimental is an issue that warrants further investigation. Rozenal et al. (2018) argued that even though adverse and unwanted events seem to exist, it is still unclear if they affect treatment outcome. Some might even be regarded as a necessary evil, such as temporary bouts of increased anxiety during exposure exercises in Cognitive Behavior Therapy (CBT). In addition, there is also an ongoing debate on how to define and measure adverse and unwanted events occurring in treatment, with different taxonomies having

been proposed, which makes it difficult to systematically assess and report such incidents across studies (Rozenal et al., 2016).

While most of the scientific literature on negative effects deal with the issue of inflicting something on the patient, e.g., novel symptoms and deterioration, less thought has been given to the *absence* of effects. Dimidjian and Hollon (2010) were early to raise the problem with non-response in treatment, arguing that no improvement at all could potentially have restricted the patient from accessing a more effective treatment. From this perspective, a treatment without any benefits would also be seen as negative given that it may have prolonged an ongoing condition and caused unnecessary suffering, and that “it still may be costly in terms of time, expense, and other resources” (p. 24). However, they also pointed out that this has to be put in relation to the natural course of the psychiatric disorder for which one has been treated, which complicates the issue of classifying non-response. Linden (2013) defined non-response as “Lack of improvement in spite of treatment” (p. 288), suggesting that it could be regarded as negative, but at the same time emphasizing the conceptual difficulties of knowing if it is caused by a properly applied treatment or not. Determining what constitutes non-response is also a question that requires a broader theoretical and philosophical discussion about treatment outcomes. Taylor et al. (2012), for instance, described some of the standards that are currently being used for identifying non-response among patients, arguing that these are often based on arbitrary cutoffs, such as a predetermined level of change or a statistical method. There is currently no consensus on how to reliably classify patients as non-responders, with many studies employing some form of diagnostic criteria, while other rely on the change scores that exceed measurement error, i.e., the Reliable Change Index (RCI; Jacobson and Truax, 1991). In a systematic review of CBT for anxiety disorders (including 87 clinical trials and 208 response rates) by Loerinc et al. (2015), the average response rate to treatment was 49.5%. In other words, about half of the patients did not respond or deteriorated. However, they noted significant heterogeneity across studies, suggesting that the response rates differ partly because of how response and non-response are defined. Looking more closely at how this was determined in the specific clinical trials revealed that 31.3% applied the RCI, 70.7% used a clinical cutoff, and 90.9% relied on some change from baseline (of note: several response rates can be used simultaneously in the same clinical trial, hence not adding up to 100%). Similar response rates have also been found in naturalistic settings when applying fixed benchmarks on self-report measures as cutoffs (Gyani et al., 2013; Firth et al., 2015), meaning that it is not uncommon for patients to experience a standstill in their treatment in a regular outpatient health care setting despite receiving the best available care.

During the last two decades, new ways of disseminating evidence-based treatments have been introduced and become an important addition to the regular outpatient health care setting. One of the most widespread formats is Internet-based CBT (ICBT), in which patients complete their treatment via a computer, tablet, or smartphone (Andersson, 2018). Similar to seeing someone face-to-face, reading material and homework assignments are considered essential components and introduced as one module per week. Patients then work on their problem and receive guidance and feedback from a therapist via email, corresponding to what would be discussed during a real-life session (Andersson, 2016). Presently, the efficacy of ICBT has been evaluated in close to 300 randomized controlled trials and several systematic reviews and meta-analyses, demonstrating its benefits for a large number of psychiatric as well as somatic conditions, including in naturalistic settings (Andersson et al., 2019). The results also seem to be maintained over time, with follow-ups at 3 years showing sustained improvements (Andersson et al., 2018). However, like treatments in general, ICBT is not without negative effects. Recent studies have for example shown that 5.8% of patients deteriorate (Rozenal et al., 2017), and that a large proportion report adverse and unwanted events (Rozenal et al., *in press*). Yet, in terms of non-response, few attempts have been made to specifically explore its occurrence and predictors. A notable exception is a study by Boettcher et al. (2014b), investigating negative effects of ICBT for social anxiety disorder. The results showed that the rate of non-responders on the primary outcome measure varied greatly during the treatment period, with 69.9% in mid-treatment, 32.3 at post-treatment, and 29.3% at 4-month follow-up. No attempt at analyzing predictors was however made. In general, the systematic study of non-response has been lacking in relation to ICBT (Andersson et al., 2014), which makes it unclear what factors might be responsible for its incidence and how this information could be used clinically (Hedman et al., 2014).

Considering the fact that a large proportion of all patients do not respond to treatment, the issue of finding those who are at risk of non-response is important. Still, few studies have explicitly explored if non-response can be predicted. Taylor et al. (2012) made an attempt at summarizing the scientific literature, describing three general factors that might prevent a patient from responding. First, poor homework adherence in CBT seems to be predictive of poorer treatment outcome, at least when it is evaluated using a sufficiently reliable and valid measure (Kazantzis et al., 2016). Second, high expressed emotion, i.e., residing in an environment characterized by hostility and emotional over-involvement, is also associated with poorer treatment outcome, but findings are mixed depending on diagnosis. Third, poorer treatment outcome is more likely if the patient displays greater symptom severity at baseline or suffers from a comorbid condition. However, in all of these cases, the focus of the research has been on responders and not explicitly non-responders, meaning that the conclusions are in fact being back-tracked. In addition, information on the standards for determining non-response have not always been clear or lacking completely. This makes it difficult to interpret the results and draw inferences to the study of non-response *per se*, making a more systematic approach to exploring the issue warranted.

Given the scarcity of research on non-response and its predictors the current study thus aims to investigate its occurrence and predictors. Seeing as ICBT is also becoming more and more common in the regular outpatient health care setting, and because it differs somewhat from seeing someone face-to-face (i.e., no or few physical meetings), it could be important to determine how often and why some patients do not seem to respond to this type of treatment. This was done by specifically looking at those patients who do not seem to benefit from ICBT, as determined using different criteria for determining non-response based on the RCI (Jacobson and Truax, 1991), and then applying a set of variables defined a priori in an analysis of possible predictors. In order to complete such a study, a large sample of individual patient data is however needed to ensure adequate statistical power (Oxman et al., 1995). Data from 29 clinical trials is therefore used, aggregated as part of a similar endeavor regarding deterioration rates (Rozenal et al., 2017). The data set consists of a total of 2,866 patients, including three categories of diagnoses: anxiety disorders, depression, and other (erectile dysfunction, relationship problems, and gambling disorder). The hypotheses are that non-response rates similar to those reported by Loerinc et al. (2015) will be obtained, i.e., 44.5%. In addition, it is also hypothesized that the findings by Taylor et al. (2012) will be seen in the current study, that is, symptom severity at baseline and module completion, a proxy for homework adherence, will constitute significant predictors of non-response, i.e., increasing the odds of not responding. Lastly, similar to Rozenal et al. (2017), not being in a relationship, younger age, and having a lower educational level are also hypothesized to be associated with increased odds of non-response.

MATERIALS AND METHODS

Individual Patient Data Meta-Analysis

To explore the rates and predictors of non-response, individual-level data from many patients are required. The current study consequently conducted an individual patient data meta-analysis, which is a powerful approach of combining the raw scores from each patient variable across studies instead of only relying on group means and standard deviations (Simmonds et al., 2005). This makes it possible to do more sophisticated statistical analyses, particularly when trying to investigate factors that might be predictive of a certain event (Oxman et al., 1995). Similar to a meta-analysis, this can be done either by performing a systematic review or pooling together data from different sites, such as university clinics. The current study used the latter method, aggregating data from those clinical trials that have been conducted by the authors and where the raw scores of patients were possible to obtain. Data from three sites run by the authors were thus screened for eligibility; (1) patients being allocated to a treatment condition involving ICBT, guided or unguided, consisting of treatment interventions that are based on CBT, including applied relaxation and cognitive bias modification (2) meeting the criteria for a psychiatric disorder or V-codes listed in the Diagnostic and Statistical Manual of Mental Disorders, Fourth or Fifth Edition (American Psychiatric Association, 2000, 2013) (3) receiving

ICBT that lasted for at least 2 weeks or two modules, and (4) completing a validated primary outcome measure assessing the patients' level of distress, for instance, for social anxiety disorder, this involved the Liebowitz Social Anxiety Scale—Self-Report (LSAS-SR; Liebowitz, 1987). Clinical trials not included in the current study were characterized by treatment conditions other than ICBT, namely, bibliotherapy with telephone support, or treatments that are not theoretically linked to CBT, such as, psychodynamic psychotherapy and interpersonal psychotherapy. A limitation of using this method is of course that it is not possible to assess the risk of bias, such as when implementing a systematic review (Simmonds et al., 2005). However, this allowed the retrieval of a majority of all clinical trials of ICBT that have been executed in Sweden, meaning that it should be representative of how it is being administered on a national level at both university clinics and in a regular outpatient health care setting, i.e., screening patients by diagnostic interviews and distributing validated outcome measures, consistent procedures for guidance by therapists, and similar distribution of treatment content.

Once the clinical trials were selected, the raw scores from each patient were put into the same data matrix and coded for consistency, e.g., sick leave (1 = yes, 0 = no) (i.e., in Sweden, receiving disability checks when being absent from work during a time period of at least 2 weeks up to 1 year due to a medical or psychiatric condition). This includes; name of the clinical trial, treatment condition, and including all available sociodemographic variables, outcome measures (primary and additional), ratings of satisfaction, and credibility, previous use of any type of psychological treatment and previous or ongoing use of psychotropic medication, sick leave, number of completed modules and time spent per week on the treatment interventions. To enable as many comparisons as possible in the statistical analysis, given that clinical trials sometimes used different coding schemes, sociodemographic variables had to be collapsed. For instance, only single/relationship were retained in terms of civil status, while the highest attained educational level was restricted to fewer but more coherent categories. Similarly, diagnoses were re-categorized to balance out their proportions: (1) anxiety disorders, and (2) depression and other (erectile dysfunction, relationship problems, and gambling disorder). Meanwhile, those numbers among the raw scores that were unclear, i.e., when information about a nominal variable was missing, published and unpublished manuscripts were obtained and checked so that the data matrix was coded in accordance with the clinical trials. However, it should be noted that the coding schedules for some of the original datasets were impossible to retrieve, whereby a few cells remained blank. For an overview of the patients' sociodemographic variables and the amount of missing data, see **Table 1**.

Statistical Analysis

Given the lack of consensus on how to define and determine non-response in treatment (Taylor et al., 2012), the RCI was chosen based on its widespread use and recognition in the scientific literature for assessing reliable change (Jacobson and Truax, 1991). This was calculated by taking the change score on

a clinical trial's primary outcome measure for a specific patient and dividing it by the standard error of difference (Speer, 1992), i.e., $SE_{diff} = SD_1 \sqrt{2} \sqrt{1-r}$, where SD_1 corresponds to the standard deviation of a condition at pre-treatment, and r is the reliability estimate (Evans et al., 1998). This calculation also takes into account possible regression to the mean effects and is often referred to as the Edwards and Nunally-method (Speer, 1992). According to Bauer et al. (2004), different ways of calculating the RCI yields similar estimates, but here Speer (1992) was chosen given that it was used in the study of deterioration by Rozenal et al. (2017). The RCI was then worked out separately for the primary outcome measure for every clinical trial and using their respective test-retest reliability rather than internal consistency (see **Table 2**), in line with the recommendations by Edwards et al. (1978). Essentially, the RCI sets the boundaries for which a change score can be deemed reliable, meaning that it would be unlikely ($p = 0.05$), without a true change actually occurring. For example, considering the first clinical trial in the current study, IMÅ, a change score of ± 10.13 is considered reliable on the Beck Anxiety Inventory (Beck et al., 1988). A change score that does not exceed ± 10.13 would then be deemed as non-response. Hence, using the RCI in this way, the change scores for the primary outcome measure in each clinical trial and for each patient was used to classify non-responders, which were dummy coded into a nominal variable (1 = yes, 0 = no). However, it should be noted that a RCI is usually calculated on the basis of a standard deviation unit of change equal to $z = 1.96$. Wise (2004) argued that this is a relatively conservative estimate, at least for investigating improvement and deterioration, proposing reliable change indexes that represents different confidence levels, i.e., $z = 1.28$ for a moderate change and $z = 0.84$ for a minor change. Although affecting the probability of rejecting the null hypothesis, $p = 0.10$ and 0.20 , this could be useful for detecting less frequently occurring events, such as deterioration, or to make the boundaries of the RCI narrower, as in non-response (i.e., a smaller change score would be required to be classified as a responder, consequently affecting the non-response rate). Again, using the clinical trial IMÅ as an example, a change score of ± 6.62 is regarded as a reliable change for $z = 1.28$, and 4.94 for $z = 0.84$. In the current study, the non-response rates for each clinical trial and the total estimates are presented for each of the reliable change indexes in order to facilitate a comparison, while only $z = 1.96$ is applied for analyzing possible predictors as it should increase power. All of the non-response rates are based on data for patients receiving treatment and not some form of control condition.

To investigate possible predictors, binomial logistic regression was applied with the dichotomized coding of non-response (1 = yes, 0 = no) used as the dependent variable. All predictors were entered into the model in one single block as independent variables, as no prior evidence exist with regard to building the model. However, in terms of choosing what variables to enter, theoretical assumptions or empirical findings were used as guidance to avoid the risk of finding spurious associations and restrict the type-I-error rate (Stewart and Tierney, 2002). Hence, the same variables used for investigating the predictors of deterioration were implemented (Rozenal et al., 2017): (1)

TABLE 1 | Sociodemographics variables of patients included in the individual patient data meta-analysis.

Variable	Treatment (n = 2,118)	Control (n = 748)	Full sample (n = 2,866)	Missing data
Gender: n (% female)	1299 (62.1)	501 (67)	1800 (63.4) ^a	27 (0.9) ^b
Age (years): M (SD)	38 (12.5)	40.6 (13.2)	38.7 (12.8)	29 (1)
Civil status: n (%)				744 (27)
Single	497 (33.5)	171 (28.1)	668 (31.9)	
Relationship	986 (66.5)	438 (71.9)	1424 (68.1)	
Children: n (% yes)	554 (53.4)	226 (59)	780 (55)	1,446 (50.5)
Cohabitant: n (% yes)	306 (66.5)	48 (69.6)	354 (12.4)	2,337 (81.5)
Highest educational level: n (%)				1,099 (38.3)
Elementary school	53 (4.5)	33 (5.7)	86 (4.9)	
High school/college	361 (30.5)	169 (28.9)	530 (30)	
University	757 (64)	380 (65.1)	1137 (64.3)	
Postgraduate	12 (1)	2 (0.3)	14 (0.8)	
Employment: n (%)				1,968 (68.7)
Unemployed	74 (10.8)	19 (9)	93 (10.4)	
Student	99 (14.4)	39 (18.6)	138 (15.4)	
Employed	469 (68.2)	138 (65.7)	607 (67.6)	
Other	13 (1.9)	11 (5.2)	24 (2.7)	
Retired	33 (4.8)	3 (1.4)	36 (4)	
Primary diagnosis: n (%)				88 (3.1)
Anxiety disorders	1,148 (55.8)	533 (74.1)	1681 (60.5)	
Generalized anxiety disorder	141 (6.8)	138 (19.2)	279 (10)	
Social anxiety disorder	708 (34.4)	257 (35.7)	965 (34.7)	
Anxiety disorder NOS	11 (0.5)	20 (2.8)	31 (1.1)	
Panic disorder (with/without agoraphobia)	86 (4.2)	30 (4.2)	116 (4.2)	
Posttraumatic stress disorder	32 (1.6)	32 (4.5)	64 (2.3)	
Anxiety disorder (with/without depression)	117 (5.7)	56 (7.8)	173 (6.2)	
Specific phobia	53 (2.6)	0 (0)	53 (1.9)	
Depression (with/without dysthymia)	475 (23.1)	69 (9.6)	544 (19.6)	
Other	436 (21.2)	117 (16.3)	553 (19.9)	
Erectile dysfunction	39 (1.9)	39 (5.4)	78 (2.8)	
Relationship problems	80 (3.9)	78 (10.8)	158 (5.7)	
Gambling disorder	317 (15.4)	0 (0)	317 (11.4)	
Sick leave: n (% yes)	42 (5.5)	25 (7.6)	67 (6.1)	1,768 (61.7)
Previous psychological treatment: n (% yes)	575 (54.1)	214 (56.5)	789 (54.7)	1,424 (49.7)
Previous or ongoing psychotropic medication: n (% yes)	366 (31.7)	156 (33.1)	522 (32.1)	1,239 (43.2)
Satisfaction with treatment: M (SD) ^c	2.9 (1)	n.a.	2.9 (1)	1,867 (88.2) ^e
Treatment credibility: M (SD) ^d	7 (2.4)	n.a.	7 (2.4)	1,535 (72.5) ^e
Modules completed: M (SD) ^f	6.5 (1.3)	n.a.	6.5 (1.3)	1,194 (56.4) ^e
Time per week: M (SD) ^g	3.6 (3.1)	n.a.	3.6 (3.1)	1,722 (81.3) ^e

n.a., not applicable; NOS, not otherwise specified.

^aValid percent, i.e., percent of available data, excluding missing data.

^bPercent, i.e., percent of complete dataset, including missing data.

^cSelf-rated 0–5.

^dSelf-rated 0–10.

^eBased on patients receiving treatment.

^fWeighted mean and standard deviation.

^gNumber of hours per week.

symptom severity at baseline, (2) civil status, (3) age, (4) sick leave, (5) previous psychological treatment, (6) previous or ongoing psychotropic medication, (7) educational level, and (8) diagnosis. Two *post-hoc* and explorative variables were also entered: (9) gender, and (10) module completion. Both symptom severity at baseline and module completion, a proxy

for homework adherence, have been put forward as predictors for non-response (Taylor et al., 2012). Meanwhile, albeit not specifically linked to non-response, male gender, lower age, and lower educational level have previously been shown to predict dropout in ICBT (Christensen et al., 2009; Waller and Gilbody, 2009; Karyotaki et al., 2015).

TABLE 2 | Test-retest reliabilities used for calculating the RCI.

Primary outcome	Test-retest reliability	Time period	Population	References
Beck Anxiety Inventory	$r = 0.81$	2 weeks	Normal	Saemundsson et al., 2011
Liebowitz Social Anxiety Scale—Self-Report	$r = 0.93$	8 weeks	Normal	Heeren et al., 2012
Panic Disorder Severity Scale—Self-Report	$\rho = 0.94$	2 days	Patient	Lee et al., 2009
Patient Health Questionnaire—9 Items	$r = 0.94$	2 weeks	Patient	Zuithoff et al., 2010
International Index of Erectile Functioning—5 Items ^a	$r = 0.84$	4 weeks	Patient	Rosen et al., 1997
Beck Depression Inventory	$r = 0.77$	b	Normal	Beck and Steer, 1996
Impact of Event Scale—Revised	$r = 0.89-0.94^c$ $M = 0.92$	6 months	Patient	Sundin and Horowitz, 2002
Generalized Anxiety Disorder—7 Items	ICC = 0.83	1 week	Patient	Spitzer et al., 2006
Penn State Worry Questionnaire	$r = 0.84$	3 weeks	Normal	Pallesen et al., 2006
Body Sensations Questionnaire	$r = 0.89$	3 months	Patient	Arrindell, 1993
Dyadic Adjustment Scale ^a	$r = 0.87$	2 weeks	Patient	Carey et al., 1993
Snake Anxiety Questionnaire	$r = 0.78$	1 month	Normal	Klorman et al., 1974
Montgomery-Åsberg Depression Rating Scale—Self-Report	ICC = 0.78	1 week	Patient	Fantino and Moore, 2009
Spider Phobia Questionnaire	$r = 0.94$	3 weeks	Normal	Muris and Merckelbach, 1996
The NORC Diagnostic Screen for Gambling Problems	$r = 0.98-0.99^d$ $M = 0.98$	1 week	Patient	Gerstein et al., 1999

RCI, reliable change index; NORC, a National Organization for Research at the University of Chicago.

^aReversed scales, higher scores indicate less problems.

^bInformation regarding the time period was unavailable.

^cSeparate estimates for the two subscales.

^dLifetime test statistic and past year test statistic.

Predictors with a $p < 0.05$ were regarded as significant and presented as Odds Ratios (OR) with their respective 95% Confidence Intervals (CI), reflecting an increase or decrease in odds of non-response in relation to a reference category. For instance, for dichotomous predictors such as sick leave, the OR reflects the adjustment in odds of non-response when the patient goes from not being on sick leave (no) to being on sick leave (yes). For the three predictors that were on continuous scales, that is, symptom severity at baseline, age, and module completion, the OR represents an increase of one standard deviation above their respective mean, i.e., these variables were standardized and centered within each clinical trial. All statistical analyses were performed using jamovi version 0.9.2.9 (Jamovi project, 2018), and on a complete case basis given that it is unclear how missing data should be treated when investigating non-response.

Ethical Considerations

The data in the current study were aggregated from several clinical trials, all with written informed consent, and all having received ethical approval from the Regional Ethical Review Board at their respective study location (please refer to the original articles for more information). The data included only the raw scores from various patient variables and no sensitive or qualitative information. Moreover, all patients were given an automatically assigned identification code in each clinical trial, e.g., abcd1234, making it impossible to identify a particular individual. In terms of the ethical issue related to the assessment of non-response in ICBT, the current study used only the raw scores from already completed clinical trials, making it impossible to, in hindsight, detect and help patients that may not have benefitted from treatment. However, because the aim of the current study is to explore the occurrence and possible

predictors of non-response, future clinical trials may be better able to monitor and assist those patients who are not responding.

RESULTS

Study Characteristics

Data from 29 clinical trials were reviewed according to predefined inclusion and exclusion criteria and deemed eligible for the current study. Raw scores from all patients were then entered into the data matrix. In total, 2,118 (73.9%) had received treatment (ICBT). The following diagnoses were included (clinical trials, k): social anxiety disorder ($k = 9$), depression (with/without dysthymia; $k = 5$), generalized anxiety disorder ($k = 3$), anxiety disorder (with/without depression; $k = 3$), mixed anxiety disorders (e.g., panic disorder as well as social anxiety disorder; $k = 2$), specific phobia ($k = 2$), post-traumatic stress disorder ($k = 1$), panic disorder (with/without agoraphobia; $k = 1$), gambling disorder ($k = 1$), erectile dysfunction ($k = 1$), and relationship problems ($k = 1$). In terms of recruitment, self-referrals from the general population were most common, 27 clinical trials, but one was conducted in primary care, and another at a university clinic. With regard to screening interviews, the Structured Clinical Interview for DSM-IV-Axis I Disorders (First et al., 1997), was mostly applied, followed by four clinical trials that implemented either the MINI-International Neuropsychiatric Interview (Sheehan et al., 1998), or a diagnosis-specific instrument, e.g., Clinician-Administered PTSD Scale (Blake et al., 1995). The length of treatment ranged from four to 10 modules ($M = 8.28$; $SD = 1.36$), 4–12 weeks ($M = 8.45$; $SD = 1.66$), and two to 10 sessions ($M = 5.40$; $SD = 3.58$), with specific phobia being shortest, while various anxiety disorders and relationship problems were the longest. The total amount of

missing data for the primary outcome measures at post-treatment was 12.9%. For a complete overview of the clinical trials, please refer to **Table 3**.

Non-response Rates

Of the 2,118 patients, 567 (26.8%) were classified as non-responders when using a RCI of $z = 1.96$. In comparison, the numbers were a bit lower, 356 (16.8%) for $z = 1.28$, and a mere 239 (11.3%) for $z = 0.84$, indicating that the non-response rates vary depending on what reliable change indexes are being employed, each step being statistically significant, $\chi^2_{(2)} = 64.89$, $p < 0.05$, and $\chi^2_{(2)} = 27.57$, $p < 0.05$. The lowest rates of non-response ($z = 1.96$) can be found in clinical trials for gambling disorder (3.5%), specific phobia for snakes (7.6%), and depression (10.9%). Meanwhile, the highest rates were obtained in clinical trials on erectile dysfunction (74.4%), and anxiety disorders (with/without comorbid depression; 58.8 and 56.6%, respectively). See **Table 3** for an outline of the non-response rates in each clinical trial, sorted according the respective reliable change indexes.

Predictors of Non-response

A binomial logistic regression was performed with the predefined variables entered as predictors for non-response. The results can be seen in **Table 4**, together with their respective OR and 95% CI. Overall, the output seems to suggest that patients receiving treatment had increased odds of non-response if they had higher symptom severity on the primary outcome measure at baseline. Similarly, there were increased odds for not responding in treatment when having an anxiety disorder as compared to depression and other (erectile dysfunction, relationship problems, and gambling disorder), and if the patient was of male gender. None of the other variables were predictive of non-response.

DISCUSSION

The current study examined the occurrence of non-response in clinical trials of ICBT for three categories of problems including anxiety disorders, depression, and other (erectile dysfunction, relationship problems, and gambling disorder). In total, 2,118 patients in 29 clinical trials received treatment and were analyzed, indicating that 567 (26.8%) were classified as non-responders when using a RCI of $z = 1.96$, but fewer when implementing a narrower criterion, 356 (16.8%) for 1.28, and 239 (11.3%) for 0.84. This goes against the initial hypothesis of finding a similar estimate as the systematic review of CBT for anxiety disorders by Loerinc et al. (2015), which found an average response rate of 44.5%, indicating that non-response could be less frequent in ICBT. However, concluding that non-response is more common in CBT is highly speculative given that such numbers may not be possible to back-track, i.e., the opposite of response also includes patients who deteriorate. Thus, it would be more correct to compare it to attempts at determining non-response more directly. For example, Gyani et al. (2013) demonstrated that 29.0% did not respond among 19,395 patients receiving treatment within Improving Access to Psychological Therapies (IAPT) in the United Kingdom, with a majority

undergoing CBT. Similarly, Firth et al. (2015) analyzed 6,111 patients from IAPT using the same method, demonstrating that 32–36% were classified as non-responders. Hence, at least according to these estimates, the rate obtained in the current study on ICBT closely resemble those for treatments delivered face-to-face, at least when using a RCI of $z = 1.96$. However, in these two cases a composite measure of non-response was in fact used, incorporating both the Patient Health Questionnaire-9 Items (PHQ-9; Löwe et al., 2004) and the Generalized Anxiety Disorder Assessment-7 Items (GAD-7; Spitzer et al., 2006). In addition, they also applied a predetermined cutoff for distinguishing responders from non-responders, which is quite different from using the RCI as it only sets one boundary, i.e., determining non-response based on having a treatment outcome *above* a certain threshold as compared to a change score *within* a particular range. In comparison, Hansen et al. (2002) used the RCI for the Outcome Questionnaire-45 (OQ-45; Lambert et al., 1996), i.e., “no change, meaning a patient’s OQ-45 score had not changed reliably in any direction over the course of therapy” (p. 337), having a non-response rate of 56.8%. Meanwhile, Mechler and Holmqvist (2016) used the RCI in relation to the Clinical Outcomes in Routine Evaluation—Outcome Measure (Evans et al., 2002), with non-response rates being 61.2–66.6% (the range in the latter depending on whether patients were in primary care or a psychiatric outpatient unit). The number of patients not responding to treatment thus seems to vary greatly depending on how this is being classified, making it difficult to draw any definite conclusions on what estimates may be more accurate. This is especially true when different studies use different categories of treatment outcome, such as when improvement is also divided into improved and recovered, thereby obfuscating the results and making direct comparisons more complicated. In addition, it is important to keep in mind what population was explored. Patients in naturalistic settings may differ from those in clinical trials, where inclusion and exclusion criteria may prevent the most severe patients from being included, hence the much higher rates. The numbers from the current study should thus be interpreted cautiously and perhaps only be compared to patients who receive treatment in a tightly controlled research setting where the internal validity is increased and the samples highly selected.

As for ICBT more specifically, comparing non-response rates is difficult. Systematic reviews have not yet explicitly investigated the issue and clinical trials do not generally determine non-responders as a separate categorical outcome. However, a few exceptions exist. Boettcher et al. (2014b) found that 32.2% of the patients receiving CBT via the Internet for social anxiety disorder did not respond when analyzing the primary outcome measure and using the RCI with an intention-to-treat principle. Likewise, Probst et al. (2018) showed that in a treatment for tinnitus distress, 20.4% could be identified as non-responders (27.2% if using an intent-to-treat analysis where missing data was classified as non-response), although, in this case, a predetermined cutoff was utilized. Based only on these examples, findings from the current study seem to be similar, but it would be useful if future clinical trials reported non-response rates more regularly to facilitate systematic reviews.

TABLE 3 | Characteristics and non-response rates for the clinical trials included in the individual patient data meta-analysis.

Study	Recruitment	Screening interview	Primary diagnosis	Treatment (n)	Control (n)	Modules/weeks or sessions	Primary outcome	Additional outcomes	Non-response 1.96 n (%)	Non-response 0.84 n (%)
IMÅ (Boettcher et al., 2014a)	General population	SCID-I	Panic disorder, social anxiety disorder, generalized anxiety disorder, anxiety disorder NOS	Unguided mindfulness with FAQ (42)	Wait-list with discussion forum (46)	8 modules/8 weeks	BAI	BDI, QOLI, ISI	12 (28.6%)	8 (19.0%) 5 (11.9%)
ACT Smart (Lindner et al., 2013)	General population	SCID-I	Panic disorder, social anxiety disorder	Unguided ACT (48), guided ACT (48)	Wait-list (47)	8 modules/10 weeks	LSAS-SR, PDSS-SR ^a	PHQ-9, GAD-7, QOLI	36 (37.5%)	25 (26.0%) 17 (17.7%)
ACTUA (Nyström et al., 2017)	General population	SCID-I	Depression	Guided physical activity (164), guided behavioral activation (158) ^b	n.a.	8 modules/12 weeks	PHQ-9	GAD-7, QOLI ^c	35 (10.9%)	15 (4.7%) 15 (4.7%)
ADAM (Andersson et al., 2011)	General population	Semi-structured interview, IIEF-5	Erectile dysfunction	Guided CBT (39)	Wait-list with discussion forum (39)	7 modules/7 weeks	IIEF-5	IIEF, RAS, BDI, BAI	29 (74.4%)	22 (56.4%) 18 (46.2%)
Challenger (Boettcher et al., 2018)	General population	MINI	Social anxiety disorder	Unguided CBT with smartphone application (68), unguided bibliotherapy (70)	Wait-list (69)	9 modules/6 weeks	LSAS-SR	GAD-7, PHQ-9, QOLI, BBQ, Mini-SPIN	46 (33.3%)	26 (18.8%) 19 (13.8%)
Stella (Andersson et al., 2013a)	General population	SCID-I & II	Depression	Guided CBT (33), group therapy (36), guided CBT as preferred choice (16)	n.a.	8 modules/8 weeks, 8 sessions	BDI	MADRS-S, BAI, HAM-D, QOLI,	47 (55.3%)	34 (40.0%) 18 (21.2%)
Depressionshjälpen (Carlbirg et al., 2013)	General population	SCID-I	Depression	Guided CBT (40)	Wait-list (40)	7 modules/8 weeks	BDI	MADRS-S, BAI, QOLI, WAI	13 (32.5%)	8 (20.0%) 7 (17.5%)
Tellus (Ivarsson et al., 2014)	General population	CAPS	Posttraumatic stress disorder	Guided CBT (31)	Wait-list with support (31) ^d	8 modules/8 weeks	IES-R	PDS, BDI, BAI, QOLI	6 (19.4%)	3 (9.7%) 3 (9.7%)
Klara (Vermark et al., 2010)	General population	SCID-I	Depression	Guided CBT (29), guided CBT via email (30)	Wait-list (29)	7 modules/8 weeks	BDI	MADRS-S, BAI, QOLI	8 (13.6%)	7 (11.7%) 3 (5.1%)
Oroshjälpen (Dahlin et al., 2016)	General population	SCID-I	Generalized anxiety disorder	Guided ACT (52)	Wait-list (51)	7 modules/9 weeks	GAD-7	PSWQ, GAD-Q-IV, BAI, MADRS-S, PHQ-9, QOLI	14 (26.9%)	5 (9.6%) 3 (5.8%)
Nova 1 (Carlbirg et al., 2011)	General population	SCID-I	Anxiety disorder with/without comorbid depression	Guided CBT (53)	n.a.	10 modules/10 weeks ^e	BAI	MADRS-S, CORE-OM, QOLI	30 (56.6%)	21 (39.6%) 13 (24.5%)
Nova 2 (Nordgren et al., 2014)	Primary care	SCID-I	Anxiety disorder with/without comorbid depression	Guided CBT (51)	Wait-list (50)	10 modules/10 weeks ^f	BAI	MADRS-S, CORE-OM, QOLI	30 (58.8%)	20 (39.2%) 10 (19.6%)

(Continued)

TABLE 3 | Continued

Study	Recruitment	Screening interview	Primary diagnosis	Treatment (n)	Control (n)	Modules/weeks or sessions	Primary outcome	Additional outcomes	Non-response 1.96 n (%)	Non-response 1.28 n (%)	Non-response 0.84 n (%)
Origo 1 (Almiöv et al., 2011)	General population	SCID-I	Generalized anxiety disorder	Guided CBT (44)	Wait-list (45)	8 modules/8 weeks	PSWQ	GAD-Q-IV, STAI, BAI, BDI, MADRS-S, QOLI	15 (34.1%)	10 (22.7%)	6 (13.6%)
Origo 2 (Andersson et al., 2012)	General population	SCID-I	Generalized anxiety disorder	Guided CBT (27) ^g	Wait-list (27)	8 modules/8 weeks	PSWQ	GADQ-IV, MADRS-S, BDI, BAI, STAI, QOLI	10 (37.0%)	10 (37.0%)	10 (37.0%)
Panik 2 (Carlbring et al., 2005)	General population	SCID-I, CIDI	Panic disorder	Guided CBT (25), face-to-face CBT (24)	n.a.	10 modules/10 weeks, 10 sessions	BSQ	ACO, MI, BAI, BDI, QOLI	12 (24.5%)	5 (10.2%)	1 (2.0%)
Pia (unpublished)	General population	SCID-I	Relationship problems	Guided CBT (80)	Wait-list with discussion forum (78)	10 modules/10 weeks	DAS	MSI, BDI, BAI, QOLI	40 (50.0%)	25 (31.3%)	17 (21.3%)
Progridi, (Ström et al., 2013)	General population	SCID-I	Depression	Guided CBT with physical activity (24)	Wait-list (24)	9 modules/9 weeks	MADRS-S	BDI, BAI, QOLI ^c	7 (29.2%)	3 (12.5%)	1 (4.2%)
Fobal orm (Andersson et al., 2013b)	General population	SCID-I	Specific phobia	Guided CBT (13), face-to-face CBT (13)	n.a.	4 modules/4 weeks, 2 sessions ^h	SNAQ	ADIS, FSS, BAI, BDI	2 (7.6%)	1 (3.9%)	1 (3.9%)
Sofie 13 (Boettcher et al., 2013)	General population	SCID-I	Social anxiety disorder	Guided CBT with CBM (61), Guided CBT without CBM (65)	n.a.	9 modules/9 weeks ⁱ	LSAS-SR	SIAS, SPS, MADRS-S, QOLI	26 (20.6%)	12 (9.5%)	4 (3.2%)
Sofie 9 (Kuckertz et al., 2014)	General population	SCID-I & II	Social anxiety disorder	Guided CBT (40)	Attention bias modification (39)	9 modules/9 weeks	LSAS-SR	SIAS, SPS, SPSQ, BAI, MADRS-S, QOLI	9 (22.5%)	7 (17.5%)	3 (7.5%)
Sofie 12 (Dagöö et al., 2014)	General population	SCID-I, MINI	Social anxiety disorder	Guided CBT with smartphone application (24)	n.a.	9 modules/9 weeks	LSAS-SR	SIAS, SPS, MADRS-S, QOLI	6 (25.0%)	2 (8.3%)	1 (4.2%)
Sofie 1 (Andersson et al., 2006)	General population	SCID-I	Social anxiety disorder	Guided CBT (32)	Wait-list (32)	9 modules/9 weeks	LSAS-SR	SIAS, SPS, SPSQ, BAI, MADRS-S, QOLI	7 (21.9%)	4 (12.5%)	2 (6.3%)
Sofie 2 (Carlbring et al., 2007)	General population	SCID-I	Social anxiety disorder	Guided CBT (29)	Wait-list (28)	9 modules/9 weeks	LSAS-SR	SIAS, SPS, SPSQ, BAI, MADRS-S, QOLI	10 (34.5%)	6 (20.7%)	4 (13.8%)
Sofie 3 (Tillfors et al., 2008)	Students	SCID-I	Social anxiety disorder	Guided CBT (19), guided CBT with group sessions (18)	n.a.	9 modules/9 weeks, 5 sessions	LSAS-SR	SIAS, SPS, SPSQ, BAI, MADRS-S, QOLI	16 (43.2%)	12 (32.4%)	11 (29.7%)
Sofie 4 (Furmark et al., 2009)	General population	SCID-I & II	Social anxiety disorder	Guided CBT with discussion forum (40), unguided bibliotherapy (40)	Wait-list (40)	9 modules/9 weeks	LSAS-SR	SIAS, SPS, SPSQ, BAI, MADRS-S, QOLI	34 (42.5%)	20 (25.0%)	12 (15.0%)

(Continued)

TABLE 3 | Continued

Study	Recruitment	Screening interview	Primary diagnosis	Treatment (n)	Control (n)	Modules/weeks or sessions	Primary outcome	Additional outcomes	Non-response 1.96 n (%)	Non-response 1.28 n (%)	Non-response 0.84 n (%)
Sofie 5 (Furmark et al., 2009)	General population	SCID I	Social anxiety disorder	Guided AR (29), guided CBT with discussion forum (29), unguided bibliotherapy (29), unguided bibliotherapy with discussion forum (28)	n.a.	9 modules/9 weeks	LSAS-SR	SIAS, SPS, SPSQ, BAI, MADRS-S, QOLI	37 (32.2%)	21 (18.3%)	12 (10.4%)
Elsa (in preparation)	General population	SCID-I	Anxiety disorder with/without comorbid depression	Guided CBT (33)	Wait-list with support (33)	8 modules/8 weeks ^d	BAI	MADRS-S, CORE-OM, PHQ-9, GAD-7, QOLI	16 (48.5%)	12 (36.4%)	12 (36.4%)
Fobal spindel (Andersson et al., 2009)	General population	SCID-I	Specific phobia	Guided CBT (13), face-to-face CBT (14)	n.a.	5 modules/4 weeks, 2 sessions ^h	SPQ	ADIS, FSS, BAI, BDI	3 (11.1%)	1 (3.7%)	0 (0.0%)
Gamble (Carlbirg et al., 2012)	General population	n.a.	Gambling disorder	Guided CBT with discussion forum (317)	n.a.	8 modules/8 weeks	NODS	HADS, QOLI	11 (3.5%)	11 (3.5%)	11 (3.5%)
Total non-response rates									567 (26.8%)	356 (16.8%)	239 (11.3%)

SCID-I, Structural Clinical Interview for DSM-IV Axis I Disorders; NOS, Not Otherwise Specified; FAQ, Frequently Asked Questions; BAI, Beck Anxiety Inventory; BDI, Beck Depression Inventory; QOLI, Quality of Life Inventory; ISI, Insomnia Severity Index; n.a., not applicable; ACT, Acceptance and Commitment Therapy; LSAS-SR, Liebowitz Social Anxiety Scale-Self-Report; PDSS-SR, Panic Disorder Severity Scale-Self-Report; PHQ-9, Patient Health Questionnaire-9 Items; GAD-7, Generalized Anxiety Disorder-7 Items; BA, Behavioral Activation; MADRS-S, Montgomery-Åsberg Depression Rating Scale-Self-Report; IIEF-5, International Index of Erectile Functioning-5 Items; CBT, Cognitive Behavior Therapy; IIEF, International Index of Erectile Functioning; RAS, Relationship Assessment Scale; MINI, The MINI-International Neuropsychiatric Interview; BBQ, Brunnsviken Brief Quality of Life Inventory; Mini-SPIN, Mini-Social Phobia Inventory; HAMI-D, Hamilton Rating Scale for Depression; WAI, Working Alliance Inventory; CAPS, Clinician-administered PTSD Scale for DSM-IV; IES-R, Impact of Event Scale-Revised; PDS, Posttraumatic Diagnostic Scale; PSMQ, Penn State Worry Questionnaire; GAD-Q-IV, Generalized Anxiety Disorder Questionnaire; CORE-OM, Clinical Outcome in Routine Evaluation-Outcome Measure; STAI, State-Trait Anxiety Inventory; CIDI, Composite International Diagnostic Disorders Interview Schedule; FSS, Fear Survey Schedule; CBM, Cognitive Bias Modification; SIAS, Social Interaction Anxiety Scale; SPS, Social Phobia Scale; SPSQ, Social Phobia Screening Questionnaire; AR, Applied Relaxation; SPQ, Spider Phobia Questionnaire; NODS, The NORC Diagnostic Screen for Gambling Problems.

^aSeparate analyses of deterioration were conducted for the two primary outcome measures depending on the diagnosis of the patient.
^bFour treatment conditions were included in the study, with/without treatment rationale, respectively, but are pooled in the current analysis.
^cAn additional outcome measure, International Physical Activity Questionnaire, IPAQ, was used in the study, but is not included in the current analysis.
^dPassive control with the possibility to contact the research team if needed.
^ePatients were able to choose ten out of 16 modules to be completed during ten weeks.
^fPatients were able to choose ten out of 19 modules to be completed during ten weeks.
^gAn additional treatment group, guided psychodynamic therapy, was also used in the study, but is not included in the current analysis.
^hOne brief orientation session and one session of three-hour prolonged exposure.
ⁱIn addition to two weeks of CBM.
^kAn additional treatment group, interpersonal psychotherapy, was also used in the study, but is not included in the current analysis.
^lAdditional outcome measures were included in the original study but lost in the raw data file.
^mPatients were able to complete up to eight modules selected by the therapist.

TABLE 4 | Significance level, odds ratios, and 95% Confidence Intervals (CI) for predictors of non-response.

Predictor (reference)	<i>p</i>	OR	Lower CI	Upper CI
Symptom severity at baseline (lower severity)	<0.001	2.04	1.53	2.72
Diagnosis, anxiety disorders (depression and other)	<0.001	5.75	2.92	11.32
Module completion (fewer)	0.12	1.09	0.98	1.22
Civil status (single)	0.13	0.66	0.39	1.13
Gender (female)	0.03	1.80	1.05	3.10
Previous psychological treatment (no)	0.99	1.01	0.57	1.77
Previous or ongoing psychotropic medication (no)	0.08	1.72	0.94	3.14
Sick leave (no)	0.51	0.54	0.08	3.45
Educational level (below university level)	0.49	1.20	0.71	2.04
Age (lower age)	0.73	1.00	0.98	1.02

p, *p*-value; OR, odds ratio; CI, 95% confidence interval.

The current study also looked at how the application of different reliable change indexes affected the non-response rate, demonstrating a range of 15.5% between the widest and narrowest criterion. This approach was based on the recommendations by Wise (2004), contending that it can be useful to assess treatment outcome using different confidence levels: "...would be of considerable help in more accurately identifying and studying those who are not unequivocal treatment successes but who are nonetheless improving and on their way to a positive outcome as well as those who are not responding to treatment." (p. 56). However, this approach was primarily proposed for improvement and deterioration, while it is less clear if it should be applied to non-response. According to Loerinc et al. (2015), the RCI also seems to be one of the less frequently used classifications of non-response, with only one-third of the clinical trials using it in their systematic review. The results presented here are therefore tentative and need to be replicated, but they do warrant some caution as to how non-response rates are interpreted in the scientific literature (Taylor et al., 2012). Moreover, different reliable change indexes result in different rates of non-response, but what standard deviation unit of change might be most accurate depends on theory and reliability, i.e., is almost two standard deviations too broad a measure of non-response? Looking closer at one of the clinical trials included in the current study, Sofie 1, a change score within ± 15.79 on the LSAS-SR (Liebowitz, 1987) classifies a patient as a non-responder when using a RCI of $z = 1.96$, but only 6.77 points for 0.84, thereby decreasing the non-response rate from 21.9 to 6.3%. More research is needed to explore what level is clinically meaningful, that is, when a statistically determined non-response is in fact seen as something negative by the patient. This could, for instance, include interviewing those who do not respond according to the RCI regarding their experiences of treatment, similar to the study by McElvaney and Timulak (2013) who addressed the issue of good and poor outcomes using a qualitative approach.

Lastly, the current study examined possible predictors of non-response in ICBT by entering a set of variables determined a priori into a binomial logistic regression. The results from this analysis suggest that patients with higher symptom severity on the primary outcome measure at baseline, having an anxiety disorder, and being of male gender might have higher odds of not responding in treatment. The fact that greater symptoms may be a predictor is not particularly surprising given that it implies more distress and potential comorbidity, similar to what was proposed by Taylor et al. (2012), which is also in line with the initial hypothesis. Higher symptom severity could also be a sign to extend the treatment period to achieve adequate treatment dosage for those patients who do not improve as expected (Stulz et al., 2013), which is seldom possible in clinical trials. As for anxiety disorders possibly being predictive of non-response, the evidence is less clear. No direct comparisons between diagnoses have previously been made for any treatment, making it difficult to evaluate if and why this would increase the odds for not responding. One idea is that non-response occurs more often among patients with anxiety disorders in ICBT because it is more difficult for a therapist to notice and adjust the treatment without a face-to-face contact (Bengtsson et al., 2015), such as when exposure exercises need to be tweaked to target the correct stimulus or more help is required to increase motivation. Meanwhile, treating depression via the Internet might be more straightforward for the patient and therefore less probable to result in non-response. However, these findings are among the first of its kind and need to be replicated before any definitive conclusions can be drawn. It should also be noted that the third category of diagnoses, other, only consisted of three randomized clinical trials. Still, both erectile dysfunction and relationship problems had among the highest rates of non-response in the current study (74.4 and 50%), which is similar to what was found for deterioration (Rozenal et al., 2017), but gambling disorder did on the other not display the same pattern (3.5%). Further research is thus warranted to see if certain diagnoses are more likely to predict non-response in ICBT. Finally, none of the other hypotheses were confirmed, i.e., module completion, not being in a relationship, younger age, and having a lower educational level were not associated with higher odds of non-response. However, being of male gender could constitute a potential predictor, which is in line with the results by Karyotaki et al. (2015) indicating that men tend to drop out from ICBT. Here, a possible difference in coping strategies was proposed as an explanation, where women may put in more effort in trying to overcome their distress, thereby exhibiting a better compliance in treatment. If this somehow also explains the difference in non-response between the genders in ICBT remains to be seen. Yet it could be that male patients have different expectations of what the treatment entails, resulting in poorer response and dropout when these are not met, something that would be interesting to explore in the future via interviews.

Limitations

The current study is relatively unique in that it has explicitly investigated non-response in treatment and the first using individual patient data meta-analysis. This is considered a gold

standard for examining effects above those found by using group means and standard deviations, particularly in relation to discovering potential predictors (Simmonds et al., 2005). However, there are several limitations that need to be considered when interpreting the results. First, few similar examples exist in the scientific literature, making it somewhat difficult to interpret both the rates of non-response and its predictors, especially since there exists no consensus on how to define and classify patients who do not respond. The findings should therefore be interpreted cautiously and warrant replications, although they might help inform researchers of what estimates to expect and variables to explore (Clarke, 2005). Here, a particular caution should be made with regard to the ORs that have been provided, as they may be difficult to interpret and use clinically. In essence, they represent a probability of an event, similar to how odds are used in betting, but cannot be directly translated into a risk of something occurring in the future (Davies et al., 1998). Also, using binomial logistic regression in investigating predictors poses several challenges, such as how to deal with continuous scales, multicollinearity, and the assumptions regarding the relationship between the independent and dependent variables. Second, the current study consists of data from 29 clinical trials with 2,118 patients receiving treatment (2,866 in total), but the aggregation was not based on a systematic review, which could introduce different biases (Stewart and Tierney, 2002), e.g., availability bias and reviewer bias. However, the authors went to great length to ensure that all available data was used and set up predefined inclusion and exclusion criteria as a way of tackling these issues (Rozenal et al., 2017). Nevertheless, this means that the results should be explored in additional context, particularly since the clinical trials included in the current study do not have to be representative of how ICBT is conducted in other settings. Third, the patients receiving treatment can be seen as characteristic of most examples of ICBT (Titov et al., 2010), but are nonetheless more often women, in their late thirties, and having a higher education level. However, compared to treatment face-to-face, this is not particularly uncommon either (Vessey and Howard, 1993), probably reflecting a greater tendency to seek help for mental health problems among this group. Still, it does limit the generalizability of the results, particularly in terms of finding predictors of non-response. Future research should thus include patients with a more heterogeneous sociodemographic background and who have not only been self-recruited to clinical trials. This problem is also relevant regarding the diagnoses that were analyzed. Albeit including a broad spectrum of conditions, some were over-represented, e.g., social anxiety disorder, while others were less represented or even lacking completely, e.g., post-traumatic stress disorder and obsessive-compulsive disorder. Depression and other (erectile dysfunction, relationship problems, and gambling disorder) were also re-categorized to balance out their proportions, which risks losing valuable information as to where the difference lies. Thus, it is probably premature to suggest that anxiety disorders constitute a predictor for non-response before a more comprehensive investigation has been made. Fourth, the implementation of the RCI as a way of determining non-response is not without criticism and should be seen as a major limitation. It is presently unclear whether it is the best way

to identify those patients who do not respond in treatment, even if there exist a statistical rationale for its use. Furthermore, although the current study followed the recommendations by Edwards et al. (1978) on establishing valid test-retest reliabilities from the literature to calculate the RCI, most estimates relied on relatively short time periods, e.g., 2–4 weeks. This might be more relevant for assessing deterioration or improvement, but not for non-response which may need to take into account longer time frames to determine the natural fluctuation of a diagnosis. It could also be argued that the application of a cutoff or diagnostic criterion is more clinically relevant. However, those thresholds might be more useful in relation to response than non-response, i.e., defining when a patient goes from a clinical to a non-clinical population (Jacobson and Truax, 1991). Predefined numbers, such as being above a certain score, also tend to be arbitrary (Taylor et al., 2012). Still, the use of the RCI to assess non-response needs to be validated by other means. This can for instance be performed by checking if a non-responding patient still fulfills diagnostic criteria or a clinician-rating remains unchanged, e.g., the Clinical Global Impressions Scale (Busner and Targum, 2007). Non-response should also be explored in a direct comparison to deterioration and improvement in future systematic reviews. This is due to the fact that non-response is a quite heterogeneous category that could include both those patients who fare worse and achieve some positive results, even though they are, statistically speaking, seen as non-responders. Lastly, the idea of non-response representing a negative effect is not clear and warrants further debate. Both Dimidjian and Hollon (2010) and Linden (2013) argued that it might prolong an ongoing condition and prevent the patient from seeking a more helpful treatment, but that it is also important to consider the normal fluctuations of many diagnoses. In most cases, lack of improvement would probably be regarded as a failure, at least by a clinician. On the other hand, with regard to more serious conditions, lack of improvement may not necessarily be equated with something detrimental for the patient, but rather a perfectly reasonable result, i.e., remaining at a certain level of functioning in chronic pain. Also, as discussed by Linden (2013), non-response does not have to be linked to treatment, but rather other circumstances that occur simultaneously. In sum, regarding non-response as a negative effect clearly needs a discussion that considers not only the approach to classifying patients as non-responders, but also a broader theoretical and philosophical perspective of treatment outcome.

CONCLUSIONS

Among 2,118 patients in 29 clinical trials receiving treatment, 567 (26.8%) were identified as non-responders in ICBT when applying a RCI of $z = 1.96$. This is somewhat in line with other investigations in the scientific literature, although the lack of consensus on how to define non-response make it difficult to compare the results. Meanwhile, possible predictors were explored using variables set a priori, indicating that patients with higher symptom severity on the primary outcome measure at baseline, having an anxiety disorder, and being of male gender could potentially have higher odds of not responding in ICBT.

However, additional research is required to replicate the findings and to determine how to best classify non-response in treatment.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [PC], upon reasonable request.

AUTHOR CONTRIBUTIONS

All the authors contributed in the process of completing the current study and writing the final manuscript. AR conducted the aggregation of raw scores into a single data matrix and completed

the statistical analysis with input from GA and PC. AR drafted the first version of the text, while GA and PC provided feedback and reviewed and revised it.

ACKNOWLEDGMENTS

The authors would like to thank all the researchers, clinicians, and administrative staff who over the years have been involved in the clinical trials that were included in the individual patient data meta-analysis. All have been of tremendous support while collecting the data and clarifying those issues that arose during the coding process.

REFERENCES

- *Almlöv, J., Carlbring, P., Kallqvist, K., Paxling, B., Cuijpers, P., and Andersson, G. (2011). Therapist effects in guided internet-delivered CBT for anxiety disorders. *Behav. Cogn. Psychother.* 39, 311–322. doi: 10.1017/S135246581000069X
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5)*. Arlington, VA: American Psychiatric Association.
- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR)*. Washington, DC: American Psychiatric Association.
- *Andersson, E., Walen, C., Hallberg, J., Paxling, B., Dahlin, M., Almlöv, J., et al. (2011). A randomized controlled trial of guided internet-delivered cognitive behavioral therapy for erectile dysfunction. *J. Sex. Med.* 8, 2800–2809. doi: 10.1111/j.1743-6109.2011.02391.x
- Andersson, G. (2016). Internet-delivered psychological treatments. *Annu. Rev. Clin. Psychol.* 12, 157–179. doi: 10.1146/annurev-clinpsy-021815-093006
- Andersson, G. (2018). Internet interventions: past, present and future. *Internet Inter.* 12, 181–188. doi: 10.1016/j.invent.2018.03.008
- *Andersson, G., Carlbring, P., Holmstrom, A., Sparthar, E., Furmark, T., Nilsson-Ihrfelt, E., et al. (2006). Internet-based self-help with therapist feedback and *in vivo* group exposure for social phobia: a randomized controlled trial. *J. Consult. Clin. Psychol.* 74, 677–686. doi: 10.1037/0022-006x.74.4.677
- Andersson, G., Cuijpers, P., Carlbring, P., Riper, H., and Hedman, E. (2014). Guided internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. *World Psychiatry* 13, 288–295. doi: 10.1002/wps.20151
- *Andersson, G., Hesser, H., Veilord, A., Svedling, L., Andersson, F., Sleman, O., et al. (2013a). Randomised controlled non-inferiority trial with 3-year follow-up of internet-delivered versus face-to-face group cognitive behavioural therapy for depression. *J. Affect Disord.* 151, 986–994. doi: 10.1016/j.jad.2013.08.022
- *Andersson, G., Paxling, B., Roch-Norlund, P., Östman, G., Norgren, A., Almlöv, J., et al. (2012). Internet-based psychodynamic versus cognitive behavioral guided self-help for generalized anxiety disorder: a randomized controlled trial. *Psychother. Psychosom.* 81, 344–355. doi: 10.1159/000339371
- Andersson, G., Rozenal, A., Shafran, R., and Carlbring, P. (2018). Long-term effects of internet-supported cognitive behavior therapy. *Expert Rev. Neurother.* 18, 21–28. doi: 10.1080/14737175.2018.1400381
- Andersson, G., Titov, N., Dear, B. F., Rozenal, A., and Carlbring, P. (2019). Internet-delivered psychological treatments: from innovation to implementation. *World Psychiatry* 18, 20–28. doi: 10.1002/wps.20610
- *Andersson, G., Waara, J., Jonsson, U., Malmæus, F., Carlbring, P., and Ost, L. G. (2009). Internet-based self-help versus one-session exposure in the treatment of spider phobia: a randomized controlled trial. *Cogn. Behav. Ther.* 38, 114–120. doi: 10.1080/16506070902931326
- *Andersson, G., Waara, J., Jonsson, U., Malmæus, F., Carlbring, P., and Öst, L. G. (2013b). Internet-based exposure treatment versus one-session exposure treatment of snake phobia: a randomized controlled trial. *Cogn. Behav. Ther.* 42, 284–291. doi: 10.1080/16506073.2013.844202
- Arrindell, W. A. (1993). The fear of fear concept: Stability, retest artifact and predictive power. *Behav. Res. Ther.* 31, 139–148. doi: 10.1016/0005-7967(93)90065-3
- Bauer, S., Lambert, M. J., and Nielsen, S. L. (2004). Clinical significance methods: a comparison of statistical techniques. *J. Pers. Assess.* 82, 60–70. doi: 10.1207/s15327752jpa8201_11
- Beck, A. T., Epstein, N., Brown, G., and Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *J. Consult. Clin. Psychol.* 56, 893–897. doi: 10.1037/0022-006X.56.6.893
- Beck, A. T., and Steer, R. A. (1996). *Beck Depression Inventory*. Manual, Svensk version (Swedish version). Fagernes: Psykologiförlaget AB.
- Bengtsson, J., Nordin, S., and Carlbring, P. (2015). Therapists' experiences of conducting cognitive behavioural therapy online vis-à-vis face-to-face. *Cogn. Behav. Therapy* 44, 470–479. doi: 10.1080/16506073.2015.1053408
- Bergin, A. E. (1966). Some implications of psychotherapy research for therapeutic practice. *J. Abnorm. Psychol.* 71, 235–246. doi: 10.1037/H0023577
- Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., et al. (1995). The development of a clinician-administered PTSD scale. *J. Trauma. Stress* 8, 75–90. doi: 10.1007/BF02105408
- *Boettcher, J., Astrom, V., Pahlsson, D., Schenstrom, O., Andersson, G., and Carlbring, P. (2014a). Internet-based mindfulness treatment for anxiety disorders: a randomized controlled trial. *Behav. Ther.* 45, 241–253. doi: 10.1016/j.beth.2013.11.003
- *Boettcher, J., Hasselrot, J., Sund, E., Andersson, G., and Carlbring, P. (2013). Combining attention training with internet-based cognitive-behavioural self-help for social anxiety: a randomised controlled trial. *Cogn. Behav. Ther.* 43, 34–48. doi: 10.1080/16506073.2013.809141
- *Boettcher, J., Magnusson, K., Marklund, A., Berglund, E., Blomdahl, R., Braun, U., et al. (2018). Adding a smartphone app to Internet-based self-help for social anxiety: a randomized controlled trial. *Comput. Human Behav.* 87, 98–108. doi: 10.1016/j.chb.2018.04.052
- Boettcher, J., Rozenal, A., Andersson, G., and Carlbring, P. (2014b). Side effects in internet-based interventions for social anxiety disorder. *Internet Interv.* 1, 3–11. doi: 10.1016/j.invent.2014.02.002
- Busner, J., and Targum, S. D. (2007). The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry* 4, 28–37.
- Carey, M. P., Spector, I. P., Lantinga, L. J., and Krauss, D. J. (1993). Reliability of the Dyadic Adjustment Scale. *Psychol. Assess.* 5, 238–240. doi: 10.1037/1040-3590.5.2.238
- *Carlbring, P., Degerman, N., Jonsson, J., and Andersson, G. (2012). Internet-based treatment of pathological gambling with a three-year follow-up. *Cogn. Behav. Ther.* 41, 321–334. doi: 10.1080/16506073.2012.689323
- *Carlbring, P., Gunnarsdóttir, M., Hedensjö, L., Andersson, G., Ekselius, L., and Furmark, T. (2007). Treatment of social phobia from a distance: a randomized trial of internet delivered cognitive behavior therapy (CBT) and telephone support. *Br. J. Psychiatry* 190, 123–128. doi: 10.1192/bjp.bp.105.020107
- *Carlbring, P., Häglund, M., Luthström, A., Dahlin, M., Kadowaki, Å., Vernmark, K., et al. (2013). Internet-based behavioral activation and acceptance-based

- treatment for depression: a randomized controlled trial. *J. Affect. Disord.* 148, 331–337. doi: 10.1016/j.jad.2012.12.020
- *Carlbring, P., Maurin, L., Törngren, C., Linna, E., Eriksson, T., Sparthan, E., et al. (2011). Individually-tailored, Internet-based treatment for anxiety disorders: a randomized controlled trial. *Behav. Res. Ther.* 49, 18–24. doi: 10.1016/j.brat.2010.10.002
- *Carlbring, P., Nilsson-Ihrfelt, E., Waara, J., Kollenstam, C., Buhman, M., Kaldo, V., et al. (2005). Treatment of panic disorder: live therapy vs. self-help via the internet. *Behav. Res. Ther.* 43, 1321–1333. doi: 10.1016/j.brat.2004.10.002
- Christensen, H., Griffiths, K. M., and Farrer, L. (2009). Adherence in internet interventions for anxiety and depression: systematic review. *J. Med. Internet Res.* 11:e13. doi: 10.2196/jmir.1194
- Clarke, M. J. (2005). Individual patient data meta-analyses. *Best Pract. Res. Clin. Obstet. Gynaecol.* 19, 47–55. doi: 10.1016/j.bpobgyn.2004.10.011
- Cuijpers, P., Reijnders, M., Karyotaki, E., de Wit, L., and Ebert, D. D. (2018). Negative effects of psychotherapies for adult depression: a meta-analysis of deterioration rates. *J. Affect. Disord.* 239, 138–145. doi: 10.1016/j.jad.2018.05.050
- *Dagö, J., Asplund, R. P., Bsenko, H. A., Hjerling, S., Holmberg, A., Westh, S., et al. (2014). Cognitive behavior therapy versus interpersonal psychotherapy for social anxiety disorder delivered via smartphone and computer: a randomized controlled trial. *J. Anxiety Disord.* 28, 410–417. doi: 10.1016/j.janxdis.2014.02.003
- *Dahlin, M., Andersson, G., Magnusson, K., Johansson, T., Sjögren, J., Håkansson, A., et al. (2016). Internet-delivered acceptance-based behaviour therapy for generalized anxiety disorder: a randomized controlled trial. *Behav. Res. Ther.* 77, 86–95. doi: 10.1016/j.brat.2015.12.007
- Davies, H. T. O., Crombie, I. K., and Tavakoli, M. (1998). When can odds ratios mislead? *BMJ* 316, 989–991. doi: 10.1136/bmj.316.7136.989
- Delgado, J., de Jong, K., Lucock, M., Lutz, W., Rubel, J., Gilbody, S., et al. (2018). Feedback-informed treatment versus usual psychological treatment for depression and anxiety: a multisite, open-label, cluster randomised controlled trial. *Lancet Psychiatry* 5, 564–572. doi: 10.1016/S2215-0366(18)30162-7
- Dimidjian, S., and Hollon, S. D. (2010). How would we know if psychotherapy were harmful? *Am. Psychol.* 65, 21–33. doi: 10.1037/A0017299
- Ebert, D. D., Donkin, L., Andersson, G., Andrews, G., Berger, T., Carlbring, P., et al. (2016). Does Internet-based guided-self-help for depression cause harm? an individual participant data meta-analysis on deterioration rates and its moderators in randomized controlled trials. *Psychol. Med.* 46, 2679–2693. doi: 10.1017/S0033291716001562
- Edwards, D. W., Y Jarvis, R. M., Mueller, D. P., Zingale, H. C., and Wagman, W. J. (1978). Test taking and the stability of adjustment scales: can we assess patient deterioration? *Eval. Q.* 2, 275–291. doi: 10.1177/0193841X7800200206
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., et al. (2002). Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *Br. J. Psychiatry* 180, 51–60. doi: 10.1192/bjp.180.1.51
- Evans, C., Margison, F., and Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evid. Based Ment. Health* 1, 70–72. doi: 10.1136/ebmh.1.3.70
- Fantino, B., and Moore, N. (2009). The self-reported Montgomery-Åsberg Depression rating scale is a useful evaluative tool in major depressive disorder. *BMC Psychiatry* 9:26. doi: 10.1186/1471-244X-9-26
- First, M. B., Gibbon, M., Spitzer, R. L., and Williams, J. B. (1997). *Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)*. Washington, DC: American Psychiatric Press.
- Firth, N., Barkham, M., Kelleth, S., and Saxon, D. (2015). Therapist effects and moderators of effectiveness and efficiency in psychological wellbeing practitioners: a multilevel modelling analysis. *Behav. Res. Ther.* 69, 54–62. doi: 10.1016/j.brat.2015.04.001
- *Furmark, T., Carlbring, P., Hedman, E., Sonnenstein, A., Clevberger, P., Bohman, B., et al. (2009). Bibliotherapy and Internet delivered cognitive behaviour therapy with additional guidance for social anxiety disorder: randomised controlled trial. *Br. J. Psychiatry* 195, 440–447. doi: 10.1192/bjp.bp.108.060996
- Gerstein, D. R., Volberg, R. A., Toce, M. T., Harwood, H., Johnson, R. A., Buie, T., et al. (1999). *Gambling Impact and Behavior Study: Report to the National Gambling Impact Study Commission*. Chicago, IL: NORC, at the University of Chicago.
- Gyani, A., Shafran, R., Layard, R., and Clark, D. M. (2013). Enhancing recovery rates: lessons from year one of IAPT. *Behav. Res. Ther.* 51, 597–606. doi: 10.1016/j.brat.2013.06.004
- Hansen, N. B., Lambert, M. J., and Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clin. Psychol.* 9, 329–343. doi: 10.1093/clipsy.9.3.329
- Hedman, E., El Alaoui, S., Lindefors, N., Andersson, E., Rück, C., Ghaderi, A., et al. (2014). Clinical effectiveness and cost-effectiveness of internet-vs. group-based cognitive behavior therapy for social anxiety disorder: 4-year follow-up of a randomized trial. *Behav. Res. Ther.* 59, 20–29. doi: 10.1016/j.brat.2014.05.010
- Heeren, A., Maurange, P., Rossignol, M., Vanhaelen, M., Peschard, V., Eeckhout, C., et al. (2012). Self-report version of the Liebowitz Social Anxiety Scale: psychometric properties of the French version. *Can. J. Behav. Sci.* 44, 99–107. doi: 10.1037/a0026249
- *Ivarsson, D., Blom, M., Hesser, H., Carlbring, P., Enderby, P., Nordberg, R., et al. (2014). Guided internet-delivered cognitive behaviour therapy for post-traumatic stress disorder: a randomized controlled trial. *Internet Interv.* 1, 33–40. doi: 10.1016/j.invent.2014.03.002
- Jacobson, N. S., and Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J. Consult. Clin. Psychol.* 59, 12–19. doi: 10.1037//0022-006x.59.1.12
- Jamovi project (2018). *Jamovi (Version 0.9) [Computer Software]*. Available online at: <http://www.jamovi.org>
- Karyotaki, E., Kemmeren, L., Riper, H., Twisk, J., Hoogendoorn, A., Kleiboer, A., et al. (2018). Is self-guided internet-based cognitive behavioural therapy (iCBT) harmful? an individual participant data meta-analysis. *Psychol. Med.* 48, 2456–2466. doi: 10.1017/S0033291718000648
- Karyotaki, E., Kleiboer, A., Smit, F., Turner, D. T., Pastor, A. M., Andersson, G., et al. (2015). Predictors of treatment dropout in self-guided web-based interventions for depression: an 'individual patient data' meta-analysis. *Psychol. Med.* 45, 2717–2726. doi: 10.1017/S0033291715000665
- Kazantzis, N., Whittington, C., Zelencich, L., Kyrios, M., Norton, P. J., and Hofmann, S. G. (2016). Quantity and quality of homework compliance: a meta-analysis of relations with outcome in cognitive behavior therapy. *Behav. Ther.* 47, 755–772. doi: 10.1016/j.beth.2016.05.002
- Klorman, R., Weerts, T. C., Hastings, J. E., Melamed, B. G., and Lang, P. J. (1974). Psychometric properties of some specific-fear questionnaires. *Behav. Ther.* 5, 401–409. doi: 10.1016/S0005-7894(74)80008-0
- *Kuckertz, J. M., Gildebrant, E., Liliequist, B., Karlstrom, P., Vappling, C., Bodlund, O., et al. (2014). Moderation and mediation of the effect of attention training in social anxiety disorder. *Behav. Res. Ther.* 53, 30–40. doi: 10.1016/j.brat.2013.12.003
- Lambert, M., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., et al. (1996). The reliability and validity of the Outcome Questionnaire. *Clin. Psychol. Psychother.* 3, 249–258. doi: 10.1002/(SICI)1099-0879(199612)3:4<249::AID-CPP106>3.0.CO;2-S
- Lee, E. H., Kim, J. H., and Yu, B. H. (2009). Reliability and validity of the self-report version of the Panic Disorder Severity Scale in Korea. *Depress. Anxiety* 26, E120–E123. doi: 10.1002/da.20461
- Liebowitz, M. R. (1987). Social phobia. *Mod. Probl. Pharmacopsychiatry* 22, 141–173. doi: 10.1159/000414022
- Linden, M. (2013). How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. *Clin. Psychol. Psychother.* 20, 286–296. doi: 10.1002/Cpp.1765
- *Lindner, P., Ivanova, E., Ly, K., Andersson, G., and Carlbring, P. (2013). Guided and unguided CBT for social anxiety disorder and/or panic disorder via the Internet and a smartphone application: study protocol for a randomised controlled trial. *Trials* 14:437. doi: 10.1186/1745-6215-14-437
- Loerinc, A. G., Meuret, A. E., Twohig, M. P., Rosenfield, D., Bluett, E. J., and Craske, M. G. (2015). Response rates for CBT for anxiety disorders: need for standardized criteria. *Clin. Psychol. Rev.* 42, 72–82. doi: 10.1016/j.cpr.2015.08.004
- Löwe, B., Kroenke, K., Herzog, W., and Gräfe, K. (2004). Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *J. Affect. Disord.* 81, 61–66. doi: 10.1016/S0165-0327(03)00198-8
- Mays, D. T., and Franks, C. M. (1980). Getting worse: psychotherapy or no treatment – the jury should still be out. *Prof. Psychol.* 11, 78–92. doi: 10.1037/0735-7028.11.1.78

- McElvane, J., and Timulak, L. (2013). Clients' experience of therapy and its outcomes in 'good' and 'poor' outcome psychological therapy in a primary care setting: an exploratory study. *Couns. Psychother. Res.* 13, 246–253. doi: 10.1080/14733145.2012.761258
- Mechler, J., and Holmqvist, R. (2016). Deteriorated and unchanged patients in psychological treatment in Swedish primary care and psychiatry. *Nord. J. Psychiatry* 70, 16–23. doi: 10.3109/08039488.2015.1028438
- Moritz, S., Nestoriuc, Y., Rief, W., Klein, J. P., Jelinek, L., and Peth, J., (2018). It can't hurt, right? adverse effects of psychotherapy in patients with depression. *Eur. Arch. Psychiatry Clin. Neurosci.* doi: 10.1007/s00406-018-0931-1. [Epub ahead of print].
- Muris, P., and Merckelbach, H. (1996). A comparison of two spider fear questionnaires. *J. Behav. Ther. Exp. Psychiatry* 27, 241–244. doi: 10.1016/S0005-7916(96)00022-5
- *Nordgren, L. B., Hedman, E., Etienne, J., Bodin, J., Kadowaki, A., Eriksson, S., et al. (2014). Effectiveness and cost-effectiveness of individually tailored internet-delivered cognitive behavior therapy for anxiety disorders in a primary care population: a randomized controlled trial. *Behav. Res. Ther.* 59, 1–11. doi: 10.1016/j.brat.2014.05.007
- *Nyström, M. B. T., Stenling, A., Sjöström, E., Neely, G., Lindner, P., Hassmén, P., et al. (2017). Behavioral activation versus physical activity via the internet: a randomized controlled trial. *J. Affect. Disord.* 215, 85–93. doi: 10.1016/j.jad.2017.03.018
- Oxman, A. D., Clarke, M. J., and Stewart, L. A. (1995). From science to practice: meta-analyses using individual patient data are needed. *JAMA* 274, 845–846. doi: 10.1001/jama.1995.03530100085040
- Pallesen, S., Nordhus, I. H., Carlstedt, B., Thayer, J. F., and Johnsen, T. B. (2006). A Norwegian adaptation of the Penn State Worry Questionnaire: factor structure, reliability, validity and norms. *Scand. J. Psychol.* 47, 281–291. doi: 10.1111/j.1467-9450.2006.00518.x
- Probst, T., Weise, C., Andersson, G., and Kleinstäuber, M. (2018). Differences in baseline and process variables between non-responders and responders in internet-based cognitive behavior therapy for chronic tinnitus. *Cogn. Behav. Ther.* 48, 52–64. doi: 10.1080/16506073.2018.1476582
- Rheker, J., Beisel, S., Kräling, S., and Rief, W. (2017). Rate and predictors of negative effects of psychotherapy in psychiatric and psychosomatic inpatients. *Psychiatry Res.* 254, 143–150. doi: 10.1016/j.psychres.2017.04.042
- Rosen, R. C., Riley, A., Wagner, G., Osterloh, I. H., Kirkpatrick, J., and Mishra, A. (1997). The international index of erectile function (IIEF): a multidimensional scale for assessment of erectile dysfunction. *Urology* 49, 822–830. doi: 10.1016/S0090-4295(97)00238-0
- Rozenal, A., Boettcher, J., Andersson, G., Schmidt, B., and Carlbring, P. (2015). Negative effects of Internet interventions: a qualitative content analysis of patients' experiences with treatments delivered online. *Cogn. Behav. Ther.* 44, 223–236. doi: 10.1080/16506073.2015.1008033
- Rozenal, A., Castonguay, L., Dimidjian, S., Lambert, M., Shafraan, R., Andersson, G., et al. (2018). Negative effects in psychotherapy: commentary and recommendations for future research and clinical practice. *BJPsych Open* 4, 307–312. doi: 10.1192/bjo.2018.42
- Rozenal, A., Kottorp, A., Boettcher, J., Andersson, G., and Carlbring, P. (2016). Negative effects of psychological treatments: an exploratory factor analysis of the negative effects questionnaire for monitoring and reporting adverse and unwanted events. *PLoS ONE* 11:e0157503. doi: 10.1371/journal.pone.0157503
- Rozenal, A., Kottorp, A., Forsström, D., Månsson, K., Boettcher, J., Andersson, G., et al. (inpress). The negative effects questionnaire: psychometric properties of an instrument for assessing adverse and unwanted events in psychological treatments. *Behav. Cogn. Psychother.* doi: 10.1017/S1352465819000018
- Rozenal, A., Magnusson, K., Boettcher, J., Andersson, G., and Carlbring, P. (2017). For better or worse: an individual patient data meta-analysis of deterioration among participants receiving internet-based cognitive behavior therapy. *J. Consult. Clin. Psychol.* 85, 160–177. doi: 10.1037/ccp0000158
- Saemundsson, B. R., Borsdottir, F., Kristjansdottir, H., Olason, D. B., Smari, J., and Sigurdsson, J. F. (2011). Psychometric properties of the Icelandic version of the Beck Anxiety Inventory in a clinical and a student population. *Eur. J. Psychol. Assess.* 27, 133–141. doi: 10.1027/1015-5759/a000059
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J. Clin. Psychiatry* 59, 22–33.
- Simmonds, M. C., Higgins, J. P. T., Stewart, L. A., Tierney, J. F., Clarke, M. J., and Thompson, S. G. (2005). Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. *Clin. Trials* 2, 209–217. doi: 10.1191/1740774505cn0870a
- Speer, D. C. (1992). Clinically significant change: Jacobson and Truax (1991) revisited. *J. Consult. Clin. Psychol.* 60, 402–408. doi: 10.1037/0022-006X.60.3.402
- Spitzer, R. L., Kroenke, K., Williams, J. B., and Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch. Intern. Med.* 166, 1092–1097. doi: 10.1001/archinte.166.10.1092
- Stewart, L., and Tierney, J. F. (2002). To IPD or not to IPD? advantages and disadvantages of systematic reviews using individual patient data. *Eval. Health Prof.* 25, 76–97. doi: 10.1177/0163278702025001006
- *Ström, M., Uckelstam, C.-J., Andersson, G., Hassmén, P., Umefjord, G., and Carlbring, P. (2013). Internet-delivered therapist-guided physical activity for mild to moderate depression: a randomized controlled trial. *PeerJ* 1:e178. doi: 10.7717/peerj.178
- Strupp, H. H., and Hadley, S. W. (1977). A tripartite model of mental health and therapeutic outcomes: with special reference to negative effects in psychotherapy. *Am. Psychol.* 32, 187–196. doi: 10.1037/0003-066X.32.3.187
- Stulz, N., Lutz, W., Kopta, S. M., Minami, T., and Saunders, S. M. (2013). Dose-effect relationship in routine outpatient psychotherapy: does treatment duration matter? *J. Couns. Psychol.* 60, 593–600. doi: 10.1037/a0033589
- Sundin, E. C., and Horowitz, M. J. (2002). Impact of Event Scale: psychometric properties. *Br. J. Psychiatry* 180, 205–209. doi: 10.1192/bjp.180.3.205
- Taylor, S., Abramowitz, J. S., and McKay, D. (2012). Non-adherence and non-response in the treatment of anxiety disorders. *J. Anxiety Disord.* 5, 583–589. doi: 10.1016/j.janxdis.2012.02.010
- *Tillfors, M., Carlbring, P., Furmark, T., Lewenhaupt, S., Spak, M., Eriksson, A., et al. (2008). Treating university students with social phobia and public speaking fears: internet delivered self-help with or without live group exposure sessions. *Depress. Anxiety* 25, 708–717. doi: 10.1002/da.20416
- Titov, N., Andrews, G., Kemp, A., and Robinson, E. (2010). Characteristics of adults with anxiety or depression treated at an internet clinic: comparison with a national survey and an outpatient clinic. *PLoS ONE* 5:e10885. doi: 10.1371/journal.pone.0010885
- *Vermmark, K., Lenndin, J., Bjärehed, J., Carlsson, M., Karlsson, J., Öberg, J., et al. (2010). Internet administered guided self-help versus individualized e-mail therapy: a randomized trial of two versions of CBT for major depression. *Behav. Res. Ther.* 48, 368–376. doi: 10.1016/j.brat.2010.01.005
- Vessey, J. T., and Howard, K. I. (1993). Who seeks psychotherapy? *Psychotherapy* 30, 546–553. doi: 10.1037/0033-3204.30.4.546
- Waller, R., and Gilbody, S. (2009). Barriers to the uptake of computerized cognitive behavioral therapy: a systematic review of the quantitative and qualitative evidence. *Psychol. Med.* 39, 705–712. doi: 10.1017/S0033291708004224
- Wise, E. A. (2004). Methods for analyzing psychotherapy outcomes: a review of clinical significance, reliable change, and recommendations for future directions. *J. Pers. Assess.* 82, 50–59. doi: 10.1207/S15327752jpa8201_10
- Zuithoff, N. P. A., Vergouwe, Y., Kling, M., Nazareth, I., van Wezep, M. J., Moons, K. G. M., et al. (2010). The Patient Health Questionnaire-9 for detection of major depressive disorder in primary care: consequences of current thresholds in a cross-sectional study. *BMC Fam. Pract.* 11:98. doi: 10.1186/1471-2296-11-98

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Rozenal, Andersson and Carlbring. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.