Journal of Affective Disorders 275 (2020) 94–108

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Review article

People with obsessive-compulsive disorder often remain symptomatic following psychological treatment: A clinical significance analysis of manualised psychological interventions

P.L. Fisher[a,b,*], M.G. Cherry[a,b,*], T. Stuart[a], J.W. Rigby[c], J. Temple[a,b]

[a] Department of Psychological Sciences, Institute of Psychology, Health and Society, University of Liverpool, Whelan Building, Quadrangle, Brownlow Hill, Liverpool L69 3GB, UK
[b] The Royal Liverpool and Broadgreen University Hospitals NHS Trust, Liverpool, UK
[c] St Helens and Knowsley Teaching Hospitals NHS Trust, UK

ARTICLE INFO

Keywords:
Obsessive-compulsive disorder
Psychological intervention
Individual patient data
Meta-analysis
Systematic review

ABSTRACT

Background: Previous meta-analyses conclude that efficacious psychological treatments for obsessive-compulsive disorder (OCD) exist. However, determining the efficacy of psychological treatments requires multiple forms of assessment. We conducted an individual patient data meta-analysis of randomised controlled trials (RCTs) of manualised psychological therapy for adults with OCD.

Methods: Four electronic databases were searched from their inception until July 2019. IPD were available for 24 (n = 1626) of 43 (n = 2455) eligible RCTs. Treatment efficacy was evaluated using clinical significance analyses (using standardised Jacobson methodology) and standardised mean difference within-group effect-size analyses. Outcomes were Yale-Brown Obsessive Compulsive Scale (Y-BOCS) scores at post-treatment and follow-up.

Results: At follow-up, large within-group effect sizes were found for treated patients (g = 1.45) and controls (g = 0.90). Treated patients were significantly more likely than controls to recover, but recovery rates were low; post-intervention, only 32% of treated patients and 3% of controls recovered; rising to 38% and 21% respectively at follow-up. Regardless of allocation, only 20% of patients were asymptomatic at follow-up. Individual cognitive therapy (CT) was most efficacious, followed by group CT plus exposure and response prevention. Self-help interventions were generally less efficacious than face-to-face approaches.

Limitations: Data were analysed from 24 of the 43 eligible RCTs. We were unable to consider the long-term efficacy of treatments because only two RCTs provided long-term (> 12 month) follow-up data.

Conclusion: Almost 80% of treated patients remain symptomatic. The efficacy of psychological interventions for patients with OCD must be enhanced.

1. Introduction

Obsessive-compulsive disorder (OCD) is characterised by the presence of recurrent obsessions and/or compulsions (American Psychiatric Association [APA], 2013). Obsessions are intrusive thoughts, impulses, or images that cause marked distress. Compulsions are repetitive behaviours or mental acts performed in response to obsessions to alleviate distress (APA, 2013). OCD is the fourth most common mental health disorder in the world (De Putter and Koster, 2017) and has an estimated lifetime prevalence of 2.3% in the United States (Ruscio et al., 2010). OCD reduces quality of life, impairs social functioning, and increases use of health care services (Bobes et al., 2001; Eisen et al., 2006; Hollander et al., 1998). Risk of suicide is also increased in OCD with the incidence being 10

Abbreviations: CI, confidence interval; CT, cognitive therapy; DSM, diagnostic and statistical manual of mental disorders; EMDR, eye movement desensitisation and reprocessing; ERP, exposure and response prevention; ICD, international statistical classification of diseases and related health problems; IPD, Individual patient data; IPD-MA, Individual patient data meta-analysis; ITT, Intention to treat; OCD, Obsessive-compulsive disorder; RCT, Randomised controlled trial; SMD, Standardised mean difference; WLC, Waitlist control; Y-BOCS, Yale-Brown Obsessive Compulsive Scale

* Corresponding authors at: Department of Psychological Sciences, Institute of Psychology, Health and Society, University of Liverpool, Whelan Building, Quadrangle, Brownlow Hill, Liverpool L69 3GB, UK.
E-mail addresses: peter.fisher@liverpool.ac.uk (P.L. Fisher), gcherry@liv.ac.uk (M.G. Cherry).

https://doi.org/10.1016/j.jad.2020.06.019
Received 9 March 2020; Received in revised form 16 June 2020; Accepted 22 June 2020
Available online 25 June 2020
0165-0327/

times higher in OCD patients than in matched controls (de la Cruz et al., 2017). Given the debilitating nature of OCD and patient preference for psychological over pharmacological treatment (McHugh et al., 2013), the provision of efficacious psychological treatments for OCD is essential.

Meta-analyses of randomised controlled trials (RCTs) conclude that cognitive therapy (CT), exposure and response prevention (ERP), and CT plus ERP (CT + ERP) are similarly efficacious psychological treatments for OCD, and achieve comparable outcomes when delivered in group and individual formats (Abramowitz et al., 2002; Eddy et al., 2004; Gava et al., 2007; Jónsson and Hougaard, 2009; Olatunji et al., 2013; Öst et al., 2015; Pearcy et al., 2016; Romanelli et al., 2014; Rosa-Alcázar et al., 2008; Schwartze et al., 2016; Skapinakis et al., 2016; Wootton, 2016).[1] As such, clinical practice guidelines internationally recommend ERP, with or without CT, as the first line psychological treatment for OCD; CT alone is recommended as an alternative (APA, 2007; Katzman et al., 2014; National Institute for Health and Care Excellence, 2005).

Most previous meta-analyses in OCD are based solely on effect sizes. However, determining the efficacy of psychological treatment requires outcomes to be assessed using multiple methods (Fisher and Wells, 2005). While effect sizes provide valuable group information about the relative efficacy of psychological treatments, they provide no information about the proportion of patients who recover (i.e. return to normal functioning), the optimal treatment outcome in mental health settings (Keller, 2003). To indicate the proportion of patients who recover following treatment, an evaluation of clinical significance is needed. To date, only two meta-analyses in OCD have evaluated recovery rates following psychological treatments. Eddy et al. (2004) and Öst et al. (2015) reported that 34–52% of OCD patients recovered following CT, ERP, or CT + ERP. However, operational definitions of recovery varied considerably across the RCTs with cut-offs ranging from ≤ 7 to ≤ 16 points on the Y-BOCS, which makes it difficult to reach a balanced assessment of treatment efficacy due to a lack of standardised outcome (Fisher and Wells, 2005; Steketee et al., 2019). Amalgamating different operational definitions of recovery is problematic. Furthermore, it should be recognised that many patients receiving psychological treatment improve – up to 85% in clinical practice (Foa et al., 2005) – yet it is unknown what proportion continue to experience debilitating symptoms. Moreover, others may deteriorate; an outcome that is often overlooked in mental health research (Barlow, 2010; Rozental et al., 2018), despite around 5–10% of patients in mental health settings deteriorating following psychological treatment (Barlow, 2010; Cuijpers et al., 2018; Rozental et al., 2017). Thus, when evaluating treatment efficacy, it is essential to use standardised indices of recovery, improvement and deterioration in order to be able to make useful comparisons across studies.

Evaluations of treatment outcome also require an adequate evaluation of the long-term effects of psychological treatment (i.e. ≥ 12 months after completion of psychological treatment). Three previous meta-analyses evaluated long-term effects (Öst et al., 2015; Pozza and Dettore, 2017; Schwartze et al., 2016); Pozza and Dettore (2017) and Schwartze et al. (2016) found no differences in efficacy between group and individual CBT; Öst et al. (2015) found no difference in efficacy between CBT and other treatments (psychological and pharmacological). However, Schwartze et al. (2016) and Öst et al. (2015) aggregated effects 1–12 and 1–60 months post-treatment, respectively making it difficult to determine the durability of treatment effects. Moreover, all three meta-analyses (Öst et al., 2015; Pozza and Dettore, 2017; Schwartze et al., 2016) aggregated effects across different types of CBT (i.e. CT, ERP, and CT + ERP). Thus, the differential long-term efficacy of CT, ERP and CT + ERP remains unclear. A fourth meta-analysis (Olatunji et al., 2013) provided no information regarding the length of the follow-up period.

Given the lack of time and resources currently available to healthcare systems, and the restricted access to healthcare services for OCD patients in rural areas (Baer and Minichiello, 2008; Goodwin et al., 2002; Marques et al., 2010; Olatunji et al., 2009), there is also a need to know the efficacy of interventions that are both inexpensive and easily accessible. Clinical practice guidelines in the UK, US and Canada recommend self-help psychological treatment (e.g. self-help books, self-help videos, and internet-based treatments) for OCD (APA, 2007; CPA, 2014; NICE, 2005). Three meta-analyses provide support for these recommendations (Dettore et al., 2015; Öst et al., 2015; Schwartze et al., 2016). The first (Dettore et al., 2015) concluded that self-help, telephone and video-conference CBT are more efficacious than control conditions and are just as efficacious as face to face CBT. However, this meta-analysis aggregated effects across the three different forms of treatment delivery (i.e. videoconference, telephone, and self-help) making it difficult to conclude if all delivery formats were equally efficacious. The other meta-analyses (Pearcy et al., 2016; Wootton, 2016) assessed the efficacy of self-help CBT for OCD, concluding pre to post-treatment within-group effect sizes of 0.51 and 1.36, respectively. However, similar to the limitations of meta-analyses that included long-term effects, outcomes were aggregated across different types of self-help CBT (i.e. CT, ERP, and CT + ERP). Thus, the differential efficacy of different types of self-help CBT (i.e. CT, ERP and CT + ERP) is unknown.

In order to obtain a better picture of treatment efficacy, we must overcome the limitations of previous meta-analyses. One solution is to undertake an individual patient data meta-analysis (IPD-MA). IPD-MAs are considered the 'gold standard' for meta-analyses (Stewart and Parmar, 1993; Stewart and Tierney, 2002). Instead of using group statistics from published RCTs, IPD-MA combines participant-level data from each relevant RCT into a common dataset. An IPD-MA is needed to evaluate the clinical significance of treatments using a standardised method across RCTs; a major omission from previous meta-analyses. While numerous methods for determining clinical significance exist, the method developed by Jacobson and colleagues (Jacobson et al., 1984; Jacobson and Truax, 1991) is the most widely used (Ogles et al., 2001) and has good construct validity (Lunnen and Ogles, 1998; Ronk et al., 2016). According to the Jacobson method, patients can be allocated to four categories: (a) 'recovered', if they make a statistically reliable change and move from a dysfunctional to a functional population; (b) 'improved', if they make statistically reliable change but remain part of a dysfunctional population; (c) 'unchanged', if they do not make statistically reliable change; and (d) 'deteriorated', if they make statistically reliable change for the worse.

To date, only one review in OCD (Fisher and Wells, 2005) has evaluated the clinical significance of psychological treatments using the standardised Jacobson method (although a recent mega-analysis used this approach to analyse data across eight treatment clinics; please see Steketee et al. (2019) for more details). Fisher and Wells (2005) applied the Jacobson method across five psychotherapy trials to determine the proportion of patients who recovered following CT and/or ERP. An alternative criterion of 'asymptomatic' was also applied (Y-BOCS score ≤ 7; Pallanti et al., 2002). Only 53% and 61% of patients respectively recovered following CT and ERP, and only 21% and 25% of patients were classed as asymptomatic respectively, indicating that most patients continue to experience OCD symptoms following treatment. Thus, the review by Fisher and Wells (2005) provides a different perspective on the efficacy of psychological treatments for OCD compared to previous meta-analyses and proposes that there is considerable scope for improvement. However, this review was conducted 15 years ago, included non-randomised trials, did not investigate the impact of different treatment formats, conducted a limited assessment of treatment effects over follow-up and pooled recovery rates across trials without

[1] Within-group effect sizes of 0.51–1.83, and between-group effect sizes of 0.57-1.50 in favour of psychological treatment compared to controls (active control, wait-list control (WLC), treatment as usual) and 0.18–0.80 in favour of psychological treatment compared to pharmacological treatment.

weighting them, thus yielding potentially misleading summary estimates (Bravata and Olkin, 2001). Moreover, the generalizability of the findings are questionable as only five trials were included in this review (n = 198) compared to 37 trials (n = 2414) in a recent meta-analysis in OCD (Öst et al., 2015).

1.1. Aims of present study

We conducted an IPD-MA to evaluate the efficacy of manualised psychological treatments for OCD patients. Our study had two aims:

(1) To update and improve the clinical significance review conducted by Fisher and Wells (2005) using standardised Jacobson methodology and asymptomatic criterion.
(2) To examine the differential efficacy of psychological treatments for OCD by treatment type (i.e. CT, ERP, and CT + ERP) and format (i.e. group, individual, and self-help) at post-treatment and follow-up.

2. Method

The conduct and reporting of this review adheres to the general principles recommended by the Centre for Reviews and Dissemination (2009) and the Meta-Analysis of Observational Studies in Epidemiology (Stroup et al., 2000). This review broadly follows the Preferred Reporting Items for Systematic Reviews and Meta Analyses – IPD (PRISMA-IPD) guidelines (Stewart et al., 2015).

2.1. Eligibility criteria

Eligibility criteria for the IPD-MA follow the PICOS framework (Liberati et al., 2009).

2.1.1. Participants

Adults meeting diagnostic criteria for a primary diagnosis of OCD through a structured diagnostic interview according to Diagnostic and Statistical Manual of Mental Disorders (DSM) 3rd edition revised (DSM III-R; APA, 1987), DSM IV (APA, 2000), DSM 5 (APA, 2013), or the International Statistical Classification of Diseases and Related Health Problems (ICD) equivalent (ICD-10, 1999) were included.

2.1.2. Interventions

Psychological interventions were defined as manualised treatments (i.e. RCTs referring to the use of a manual to standardise treatment) using psychological techniques (Temple et al., 2018).

2.1.3. Comparators

RCTs using either no treatment, WLC, treatment as usual, placebo control (i.e. a control for nonspecific factors), or an alternative psychological treatment were included.

2.1.4. Outcomes

Either self-report or clinician rated Y-BOCS scores (Goodman et al., 1989, 1989) were used as the primary outcome. The Y-BOCS is the 'gold standard' measure of OCD severity in RCTs (Kyrios et al., 2015). The measure has good inter-rater reliability and high internal validity (Federici et al., 2010; Goodman et al., 1989; Grabill et al., 2008; Woody et al., 1995). While low convergence has been found between the self-report and clinician rated Y-BOCS on some items, there is good construct validity between the two versions on total Y-BOCS scores (Federici et al., 2010; Rosenfeld et al., 1992; Steketee et al., 1996). Thus, to be more inclusive we included RCTs using either measure.

2.1.5. Studies

Only RCTs published in English in a peer-reviewed journal were included.

2.2. Search strategy

AMED, CINAHL Plus, Medline and PsycINFO were searched for relevant RCTs from their inception until July 2019, using the following search terms: ((obsessive-compulsive disorder) or OCD) AND (effectiveness or efficacy or versus or randomi*ed or treatment or random or compar* or RCT or combin*) AND ((cognitive therap*) or (behavio*r therap*) or (cognitive behavio* therap*) or ERP or exposure or therapy or treatment or intervention) NOT (paediatric or paediatric or youth or adolescent or child*). Methodological filters limiting search results to title and English language only were applied for each database. Reference lists of included studies and systematic reviews and meta-analyses were hand-searched for additional relevant literature.

2.3. Screening and selection

Titles and abstracts were first screened for their relevance by one reviewer (TS or JT). Next, full text copies of potentially relevant papers were examined by the same reviewer (TS or JT). At each stage, 50% of potentially relevant papers were independently assessed by a second reviewer (JWR) to assess for consistency in selection. Discrepancies were resolved through discussion with two other reviewers (PF and MGC).

2.4. Assessment of risk of bias

Risk of bias in the included studies was assessed using the Cochrane Collaboration Risk of Bias tool (Higgins and Green, 2011). Risk of bias was assessed across seven areas: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting and other biases. Risk of bias assessment was completed by TS and cross-checked for accuracy by JWR; disagreement or uncertainty was resolved through consensus or arbitration by MGC.

2.5. Data extraction

The authors of eligible RCTs were contacted and anonymised raw data on the Y-BOCS at pre-treatment, post-treatment and follow-up, plus treatment condition, age, and gender of each participant were requested. If both the self-report and clinician rated Y-BOCS were used in an RCT, we requested IPD for the version which was used as their primary outcome measure. One reviewer (TS) independently extracted group level data from published reports of eligible RCTs, including year of publication, country of origin, type of treatment and control conditions, mode of delivery of treatment and control conditions, and duration and number of sessions of treatment and control conditions. A second reviewer (MGC or JWR) independently checked the accuracy of data extraction. Disagreement or uncertainty was resolved through discussion with the wider research team until consensus was reached.

2.6. Coding of treatment categories

Psychological treatments were coded into four broad treatment categories: CT, ERP, CT + ERP, or 'other' treatment:

'CT' was defined as treatment that targeted negative automatic thoughts or cognitive beliefs using cognitive techniques. Behavioural experiments to test negative automatic thoughts and cognitive beliefs could be incorporated but prolonged ERP could not. CT was split into three further categories:

(1) *'Individual CT'* (CT treatments delivered 1:1)
(2) *'Group CT'* (CT treatments delivered to groups of patients simultaneously)
(3) *'Self-help CT'* (CT treatments delivered using the internet or self-help materials)

'ERP' was defined as treatment that used exposure and response prevention but not cognitive techniques. ERP was split into three further categories:

(1) *'Individual ERP'* (ERP treatments delivered 1:1)
(2) *'Group ERP'* (ERP treatments delivered to groups of patients simultaneously)
(3) *'Self-help ERP'* (ERP treatments delivered using the internet or self-help materials)

'CT + ERP' was defined as treatment that used both cognitive techniques as defined for CT above and exposure and response prevention techniques as defined for ERP above. CT + ERP was split into three further categories:

1 *'Individual CT + ERP'* (CT + ERP treatments delivered 1:1)
2 *'Group 'CT + ERP'* (CT + ERP treatments delivered to groups of patients simultaneously)
3 *'Self-help CT + ERP'* (CT + ERP treatments delivered using the internet or self-help materials)

'Other' treatments were defined as psychological interventions that did not fit a defined category.

2.7. Coding of control categories

There were two categories of control condition:

(1) *'Placebo control'* (participants received an intervention to control for nonspecific factors)
(2) *'WLC'* (participants received no treatment)

Consensus regarding coding of treatment and control conditions was achieved through discussion among JT, MGC and PF.

2.8. General analysis strategy

We analysed outcomes at two time periods: post-treatment, defined as the earliest assessment point ≤4 weeks after treatment ended; and follow-up, defined as 3–6 months after treatment ended, giving preference to the assessment point closest to 6 months (we had planned to evaluate treatment effects ≥12 months after treatment ended but only two RCTs provided such data). Treatment efficacy was evaluated using two methods: clinical significance analysis and IPD within-group effect size analysis; both of which were conducted using a standard two-stage IPD approach (Riley et al., 2010).The two methods had the following elements:

2.8.1. Clinical significance analysis

(1) Calculation of recovery rates using Jacobson's first and second criteria, and calculation of improvement and deterioration rates using Jacobson's second criterion only.
(2) Calculation of asymptomatic rates using the asymptomatic criterion.

2.8.2. Within-group effect size analysis

(1) Calculation of standardised mean difference (SMD) within-group effect sizes.

2.9. Statistical analysis

2.9.1. Preliminary analysis

Not all eligible RCTs provided IPD. Therefore, we compared pooled pre-post SMD within-group effect sizes of RCTs providing IPD with those not providing IPD to assess whether treatment effects differed. As many studies did not include a control group, we did not calculate between-group effect sizes comparing treatment vs. control conditions. For the RCTs not providing IPD, effect sizes and 95% confidence intervals (CIs) were calculated using the data available from published reports.

2.9.2. Clinical significance analysis

Jacobson's clinical significance criteria. A cut-off point to determine whether a patient was more likely to be drawn from a functional or dysfunctional population was calculated for the Y-BOCS. To determine this cut-off, criterion a), according to classifications by Jacobson and colleagues (Jacobson et al., 1984; Jacobson and Truax, 1991) was used i.e. patients' post-treatment score falls outside the range of the dysfunctional population, defined as falling at least two standard deviations (SDs) beyond the mean of the dysfunctional population, in the direction of functionality (a detailed description of the different criteria available to determine a cut-off can be found in Jacobson and Truax [1991]). The RCI was calculated using the formula presented in Jacobson and Truax (1991). An RCI greater than ± 1.96 is required for the change to be statistically reliable at $p < 0.05$. The data used to calculate cut-off point a) and the RCI are presented in Table 1. The cut-off point was 12, meaning that post-treatment or follow-up scores ≤ 12 fell within the functional population. The RCI was 5.29 meaning that a 10-point change in score on the Y-BOCS from pre- to post-treatment or pre-treatment to follow-up was required for reliable change (i.e. ± 1.96 × 5.29). The cut-off point and RCI calculated in the present review are the same as those used in Fisher and Wells (2005).

Asymptomatic criterion. Fisher and Wells (2005) highlighted that individuals with a score of 12 on the Y-BOCS could still meet diagnostic criteria for OCD and experience considerable distress. Consequently, an assessment of the proportion of patients who were asymptomatic following treatment was conducted. The same asymptomatic criterion used in Fisher and Wells (2005) was used in the present study, namely a score of ≤ 7 on the Y-BOCS (Pallanti et al., 2002). This score indicates that a person will have minimal OCD symptoms i.e. similar to the levels experienced by individuals without OCD. While this asymptomatic criterion has not been empirically validated, it has been used in numerous studies to reflect minimal OCD symptoms (Cottraux et al., 2001; Fisher and Wells, 2005; Jones et al., 2012; Vaccaro et al., 2010, 2014; Veale et al., 2015).

Clinical significance analysis overview. The proportion of patients classed as 'asymptomatic', 'recovered', 'improved', or 'deteriorated'

Table 1
Data used to determine cut-off point (a) and the RCI on the Y-BOCS.

Symbol	Definition	Value
*M*1	Mean of the Y-BOCS at pre-treatment for the OCD sample[a,b]	24.38
*S*1	Standard deviation of the Y- BOCS at pre-treatment for the OCD sample	6.00
*X*1	Pre-treatment Y-BOCS score of an individual	
*X*2	Post-treatment score of an individual	
*r*xx	Reliability of the Y-BOCS[c]	0.61
SE	Standard error of measurement for the Y-BOCS	3.75
*S*diff	Standard error of difference between the two test scores	5.30

Note. M1 = pre-treatment mean; SD1 = pre-treatment standard deviation; Sdiff = pre-treatment standard error of difference; Rxx = test-retest reliability

[a] All pre-treatment scores (n = 1626) from 24 studies – we included pre-treatment scores on the Y-BOCS from one additional treatment group from Nakatani et al. (2005) for which we received IPD (n = 10) but excluded from the IPD-MA because it did not involve a psychological intervention (i.e. it was a pharmacological intervention)

[b] 1% of all participants could not improve or recovery as their Y-BOCS score at pre-treatment was below 10.

[c] Test-retest reliability (Woody et al., 1995); Y-BOCS = Yale–Brown Obsessive Compulsive Scale.

following treatment, and 95% CIs, were calculated for all treatment and control categories. Overall weighted proportions and weighted proportions for each treatment and control category were calculated at post-treatment and follow-up. A weighted pooled event rate was calculated across RCTs using a random effects model (Borenstein et al., 2009). Event rates were then converted to percentages. Sensitivity analyses excluding event-rate outliers were conducted. An event rate was considered an outlier when its 95% CI was outside the 95% CI of the overall mean effect size. Subgroup analyses were used to examine the influence of different treatment categories on outcome. Firstly, the weighted proportion of different treatment categories was examined. Next, Cochrane's Q test was used to explore whether outcome was moderated by treatment category. Finally, if outcome was moderated by treatment category, 95% CIs of the different treatment categories were inspected: non overlapping 95% CIs indicated significant between group differences. The same process was conducted to explore whether the control category influenced outcome. Proportions for treatment and control categories were only pooled if a minimum of two RCTs provided data. Heterogeneity between studies was also examined. The proportion of total variation that was due to heterogeneity was expressed as the I^2 statistic, with values greater than 50% indicating at least moderate heterogeneity (Higgins et al., 2003). Analyses were conducted on a per-protocol rather than intent-to-treat (ITT) basis because, at post-treatment and follow-up, we only received IPD for patients who completed post-treatment and/or follow-up assessments. Analyses were performed using Comprehensive Meta-Analysis, version 3.3.07.

2.9.3. Within-group effect size analysis

SMD within-group effects-sizes. Because many studies did not include a control group, we did not calculate between-group effect sizes comparing treatment vs. control conditions. Instead we calculated within-group effect sizes comparing pre-post and pre-follow-up change. SMD within-group effect sizes with 95% CIs were calculated using IPD by using the following formula (Cohen, 1988), where M1 is the pre-treatment mean, M2 is the post-treatment or follow-up mean, Sdifference is the standard deviation of the mean difference, and r is the correlation between the pre-treatment and post-treatment/follow-up scores.

$$SMD = \left(\frac{M1 - M2}{\text{Sdifference}}\right)\sqrt{2(1-r)}$$

SMDs were adjusted for small-sample bias using Hedges' *g* (Hedges, 1989) and were pooled across RCTs using the inverse variance random effects model (DerSimonian and Laird, 1986). Overall SMD and 95% CIs and SMD and 95% CIs for each treatment and control category were calculated at post-treatment and follow-up. All SMDs were scaled so that positive values represented a larger reduction in OCD symptomatology over time. The same techniques as in the proportion analysis were used to identify outliers, and explore whether effect sizes were influenced by different treatment or control categories. Effect size analyses were conducted on a per-protocol basis. Analyses were conducted using Comprehensive Meta-Analysis, version 3.3.07.

3. Results

3.1. Study selection

The database search retrieved 2637 citations; 22 more were identified through hand searching. After removal of duplicates, 1699 remained for screening based on title and abstract. Of these, 1592 clearly did not meet inclusion criteria. Full text articles of the remaining 107 citations were retrieved and assessed. In total, 44 papers corresponding to 43 RCTs were eligible. IPD were available for 24 (n = 1626) of the 43 (n = 2455) eligible RCTs. Fig. 1 shows the study selection process. A complete list of references of eligible RCTs can be found in Appendix A in the online supplementary material.

3.2. Study and patient characteristics

Table 2 describes study and participant characteristics for the 24 included RCTs. In total, 40 treatment and 13 control groups were included in our analyses (n = 1626). Mean Y-BOCS scores across the 40 treatment conditions ranged from 18.03 to 31.86; over half of patients had moderate OCD (Goodman et al., 1989a, 1989b). Of the 40 treatments, 18 were categorised as ERP (10 were individual, six were self-help, and two were group). Thirteen treatments were categorised as CT + ERP (six were group, four were self-help, and three were individual). Eight were categorised as CT (five were individual, two were group, and one was self-help). Only one treatment was categorised as 'other' (EMDR; eye movement desensitisation and reprocessing). The mean duration of treatment (excluding self-help treatments) was 18 h (median 18; range 5–32) over a mean of 13 sessions (median 12; range 6–20). In the self-help treatments, participants were instructed to use the self-help material for an average of 11 weeks across the different RCTs (median 12, range 8–12).

Of the 13 control groups, seven were categorised as WLC and six as 'placebo controls' (psychoeducation material on OCD, online non-directive supportive therapy, applied relaxation techniques such as deep breathing, progressive muscle relaxation, and autogenic training, structured problem solving; and stress management). The duration of 'placebo controls' was reported for 4 RCTs. Of these, the mean duration was 10.13 h (n = 4; median 10.5; range 4.5–15) over a mean of 10 sessions (n = 3; median 12; range 6–12). All post-treatment assessments took place within 4 weeks of treatment completion Of the 19 treatment and 4 control conditions reporting follow-up data, the mean follow-up assessment took place four months after treatment ended (median 3; range 3–6 months).

3.3. Risk of bias

The methodological quality of included RCTs is summarised in Table 3. Most included studies did not provide sufficient information to accurately assess risk of bias in certain areas, resulting in these areas being rated as 'unclear'.

Random sequence generation: although all studies reported randomising participants to treatment or control conditions, ten studies did not describe random sequence generation in enough detail to allow a definite judgement of risk of bias. Furthermore, most studies (n = 15; 60%) reported little to no information about allocation concealment.

Blinding: as is common in psychological research, no participants or therapists were blinded to intervention allocation. However, 10 studies provided insufficient information about study personnel to determine risk of detection bias.

Incomplete outcome data and selective reporting: All studies provided details of attrition rates, with many at low risk of attrition bias. Most studies analysed data on an ITT basis, and adequately described the way that missing data were handled.

Other sources of bias: most trials were deemed to be at low risk of other sources of bias.

3.4. Proportion of missing data

At post-treatment, the weighted proportion of missing data for treated patients ranged from 2 to 28% across the different treatment categories. This compared to 7–9% across the control group categories. The proportion of missing data could not be calculated for self-help CT or 'other' treatments as not enough trials (n = 1) provided post-treatment data. At follow-up, weighted mean proportion of missing data for treated patients ranged from to 5–34% across the different treatment categories. This compared to 16% in placebo controls. Dropout rates were not calculated for group CT, self-help CT, self-help ERP, or 'other' treatments as not enough trials (n = 1) provided follow-up data (see Table 4).

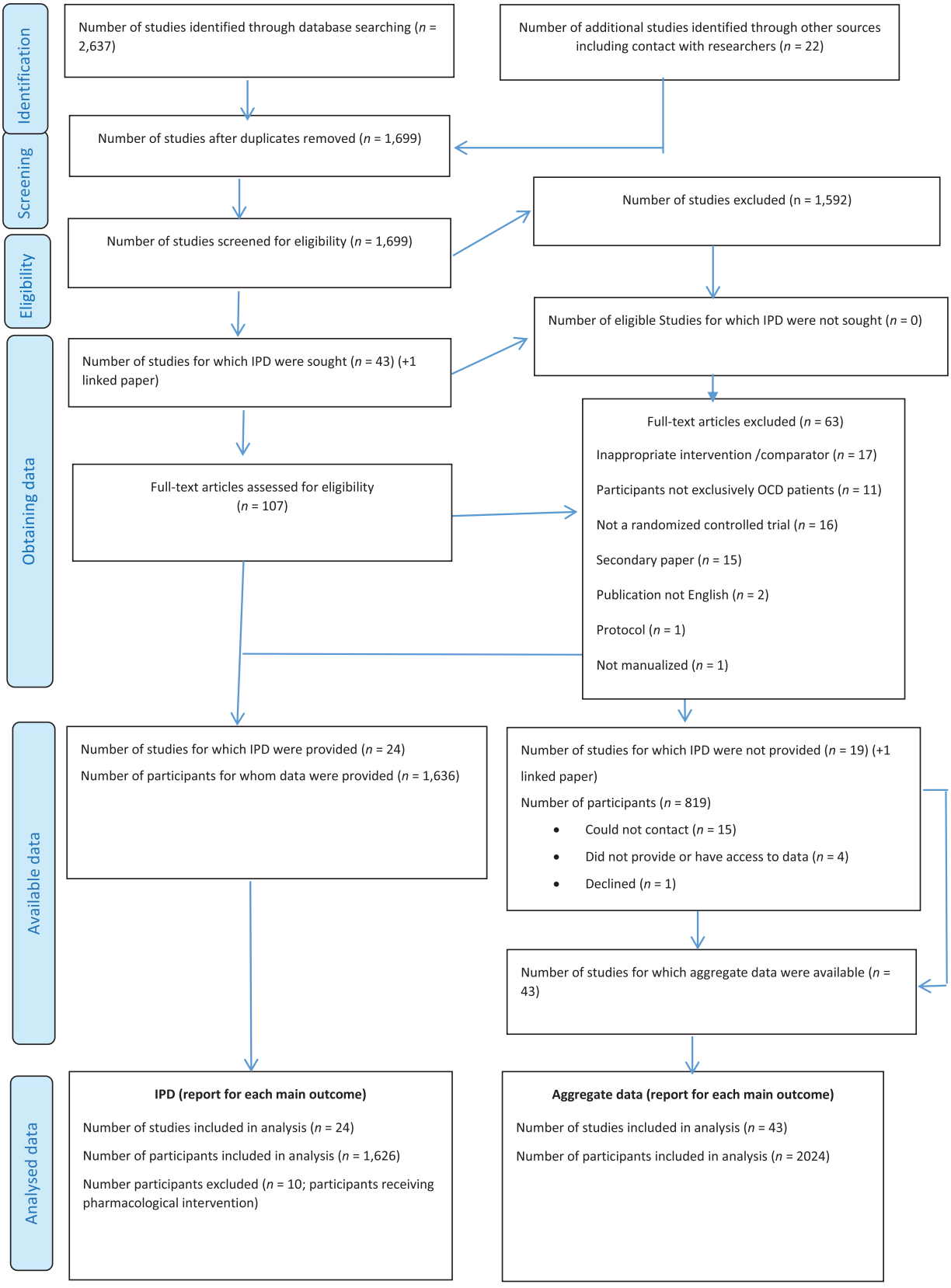

Fig. 1. Flow diagram.

Table 2
Study and patient characteristics for included trials.

Study	Location	Funding source	Condition	Sample size in report	IPD sample size[a]	Male (*n*)[a]	Female (*n*)[a]	Age, mean (*SD*)[a]	Y-BOCS pre-test mean[a]	Treatment intensity	Total hours
Anderson and Rees (2007)	Australia	Division of Health Sciences, Curtin University of Technology	Group CT + ERP	25	20	4	16	34.6 (15.89)	25.45 (7.25)	10 × 2 h sessions over 10 weeks	20
			Individual CT + ERP	21	17	6	11	32.24 (7.60)	24 (6.2)	10 × 1 h sessions over 10 weeks	10
			WLC	17	17	6	11	34.94 (9.45)	23.65 (4.73)	0 sessions	0
Andersson et al. (2012)	Sweden	The Swedish Research Council and the Swedish Society of Medicine	Self-help CT + ERP	50	49	17	32	34.02 (11.62)	21.31 (4.57)	10 modules over 10 weeks	N/R
			Placebo control[b]	51	51	17	34	36.06 (13.74)	20.8 (4.04)	10 weeks (sessions and hours N/R)	N/R
Baruah et al. (2018)	India	The Department of Biotechnology (DBT), Government of India grant (BT/PR13334/Med/30/259/2009)	Group ERP	30	29	17	12	30.17 (8.00)	24.9 (3.93)	6 × 90–120 min sessions over 3–4 weeks	9–12
			Placebo control[c]	34	28	14	14	31.32 (7.60)	25.14 (2.62)	6 × 30–45 min sessions	3–4.5
Belloch et al. (2011)	Spain and Argentina	The Spanish Ministerio de Ciencia e Innovación (SEJ2006/03893-PSIC, and PSI2010–18,340).	Individual CT	18	16	6	10	30.44 (5.70)	25.81 (4.86)	18 × 1 h sessions over 6 months	18
			Group CT	26	22	11	11	37.14 (10.48)	26.55 (4.95)	16 × 2 h sessions over 4 ½ months	32
Cordioli et al. (2003)	Brazil	The Hospital de Cl´nicas of Federal University of Rio Grande do Sul	Group CT + ERP	23	22	8	14	N/R	26.55 (4.95)	12 × 2 h sessions over 12 weeks	24
			WLC	24	23	11	12	N/R	25 (5.07)	0 sessions	0
Cottraux et al. (2001)	France	The French Ministry of Health (PHRC 95 031)	Individual CT	32	30	5	25	36.83 (9.80)	28.6 (5.14)	20 × 1 h sessions over 16 weeks	20
			Individual ERP	32	30	11	19	34.83 (11.58)	28.33 (5.02)	20 × 20 120 min sessions over 16 weeks	20
Gomes et al. (2016)	Brazil	Fundo de Incentivo a Pesquisa e Eventos/Hospital de Clínicas de Porto Alegre (FIPE-HCPA) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).	Group CT + ERP	74	49	20	29	44.69 (14.39)	27.71 (5.65)	12 × 2 h sessions over 12 weeks	24
			WLC	66	39	13	26	38.03 (13.93)	25.79 (4.81)	0 session	0
Hauschildt et al. (2016)	Germany	The German Federal Ministry of Research and Education (01GX1010).	Self-help CT	64	50	18	32	41.14 (11.23)	21.96 (6.47)	N/R	N/R
			Placebo control[b]	64	54	18	36	40.19 (9.37)	21.61 (6.63)	N/R	N/R
Herbst et al. (2014)	Germany	The German Research Foundation (KU 2754/1–1)	Self-help CT + ERP	16	16	5	11	38.19 (8.80)	20.25 (6.71)	14 sessions over 8 weeks	N/R
			WLC	18	18	7	11	33.22 (9.50)	20 (5.4)	0 sessions	0
Jaurrieta et al. (2008)	Spain	Generalitat de Catalunya (2005SGR00322)	Individual CT + ERP	19	12	8	4	34.75 (12.48)	26.83 (8.79)	20 × 45 min sessions over 20 weeks	15
			Group CT + ERP	19	16	11	5	30.19 (6.12)	24.62 (7.84)	20 × 90 min sessions over 20 weeks	30
			WLC	19	19	10	9	32.53 (10.03)	24.79 (7.27)	0 sessions	0

(*continued on next page*)

Table 2 (*continued*)

Study	Location	Funding source	Condition	Sample size in report	IPD sample size[a]	Male (*n*)[a]	Female (*n*)[a]	Age, mean (*SD*)[a]	Y-BOCS pre-test mean[a]	Treatment intensity	Total hours
Kobak et al. (2015)	USA	The National Institute of Mental Health, National Institutes of Health, and Department of Health and Human Services (R43MH090612)	Self-help ERP (ERP alone)	28	28	9	19	N/R	22.82 (3.68)	12 weeks website use	N/R
			Self-help ERP (ERP + lay coaching)	28	28	9	19	N/R	22.71 (3.97)	12 weeks website use, scheduled weekly phone calls	N/R
			Self-help ERP (ERP + therapist coaching)	31	31	14	17	N/R	21.81 (4.05)	12 weeks website use, scheduled weekly phone calls	N/R
Lindsay et al. (1997)	Australia	Not stated	Individual ERP	9	9	6	3	N/R	28.67 (4.56)	15 h over 3 weeks	15
			Placebo control[d]	9	9	0	9	N/R	24.44 (6.98)	15 h over 3 weeks	15
Lovell et al. (2006)	UK	NHS Executive North West (Research and Development Fund).	Individual ERP (via telephone)	36	35	16	19	33.49 (9.22)	24.74 (4.71)	2 × 1 h face to face sessions + 8 scheduled calls up to 30 min, weekly	2–6
			Individual ERP (face to face)	36	33	12	21	30.15 (10.17)	23.94 (5.57)	10 × 1 h sessions	10
Lovell et al. (2017)	UK	The National Institute for Health Research Health Technology Assessment programme (09–81–01)	Self-help CT + ERP (computerised)	157	106	41	65	36.64 (12.70)	24.87 (5.37)	Advised to use 6 times over 12 weeks, plus 6 × 10-min telephone calls	N/R
			Self-help CT + ERP (book)	158	119	44	75	35.29 (12.93)	25.1 (4.6)	Self-help book plus 1 × 1 h and 10 × 30 min sessions over 12 weeks	N/R
			WLC	158	121	54	67	36.76 (12.24)	12.24 (5.36)	0 sessions	0
Marsden et al. (2018)	UK	Leeds Community Healthcare NHS Trust	'Other' treatment[e]	29	15	5	10	31.33 (10.98)	24.47 (7.11)	16 sessions	N/R
			Individual ERP	26	18	6	12	35.00 (17.19)	27.17 (7.65)	16 sessions	N/R
McLean (2001)	Canada	The British Columbia Health Care Research Foundation (06 97–1; 211 95–1)	Group CT	49	31	16	15	35.87 (9.57)	21.9 (5.83)	12 × 2.5 h sessions over 12 weeks	30
			Group ERP	44	32	17	15	33.94 (10.02)	21.78 (4.56)	12 × 2.5 h sessions over 12 weeks	30
Meyer et al. (2010)	Brazil	Not stated	Group CT + ERP (CT + MI + TM)	48	47	9	38	38.57 (12.69)	30.68 (4.23)	12 × 2 h sessions weekly + 2 × 1h MI + TM sessions before group sessions	26
			Group CT + ERP (CT only)	45	43	12	31	38.30 (12.73)	31.86 (4.55)	12 × 2 h sessions weekly + 2 1 h control sessions	26
Nakatani et al. (2005)	Japan	The Ministry of Education, Culture, Sports, Science and Technology (C14570931) and the Ministry of Health, labour Welfare (14A-1)	Self-help ERP	10	10	3	7	32.50 (11.15)	29.9 (3.11)	12 × 45 min sessions weekly over 12 weeks	9
			Placebo control[f]	8	8	3	5	35.88 (8.73)	30.5 (3.67)	12 × 45 min sessions weekly over 12 weeks	9
Tolin et al. (2007)	USA	The Patrick and Catherine Weldon Donaghue Medical Research Foundation	Individual ERP	21	17	8	9	35.47 (12.99)	24.06 (4.93)	15 sessions twice weekly over 7 ½ weeks	N/R
			Self-help ERP	20	17	13	4	40.88 (13.02)	22.29 (3.74)	Instructed reading over 6 weeks	N/R
van Oppen et al. (1995)	Netherlands	Not stated	Individual CT	35	28	26	2	34.1 (10.8)	24.14 (5.12)	16 × 45 min sessions over 16 weeks	12
			Individual ERP	36	29	0	29	35.3 (10.1)	25.41 (7.01)	16 × 45 min sessions over 16 weeks	12

(*continued on next page*)

Table 2 (*continued*)

Study	Location	Funding source	Condition	Sample size in report	IPD sample size[a]	Male (*n*)[a]	Female (*n*)[a]	Age, mean (*SD*)[a]	Y-BOCS pre-test mean[a]	Treatment intensity	Total hours
Vogel et al. (2014)	Norway	The Norwegian Extra Foundation for Health and Rehabilitation (2009/3/0075).	Individual ERP	10	10	4	6	28.80 (9.25)	24.2 (4.29)	9 × 30–45 min telephone + 6 × 90 min videoconference sessions over 12 weeks	13.5 - 15.75
			Self-help ERP	10	10	5	5	29.80 (10.34)	24.1 (2.73)	12 weeks	N/R
			WLC	10	10	3	7	40.70 (11.08)	23.4 (2.8)	0 sessions	0
Vogel et al. (2004)	Norway	The National Council for Mental Health Norway	Individual ERP + CT	16	12	5	7	28.67 (7.57)	25.42 (2.94)	2 × 2 h sessions weekly over 6 weeks	24
			Individual ERP	19	15	2	13	38.87 (13.71)	23.2 (2.86)	2 × 2 h sessions weekly over 6 weeks	24
Whittal et al. (2005)	Canada	Not stated	Individual CT	34	30	8	22	35.57 (9.67)	23.5 (4.33)	50–60 min sessions over 12 weeks	10–12
			Individual ERP	37	29	14	15	34.24 (11.31)	21.66 (5.92)	50–60 min sessions over 12 weeks	10–12
Whittal et al. (2010)	Canada	The Canadian Institutes of Health Research.	Individual CT	40	37	17	20	31.57 (8.91)	18.03 (6.29)	12 × 1 h sessions over 12 weeks	12
			Placebo control[g]	33	32	20	12	31.31 (10.70)	17.28 (7.71)	12 × 1 h sessions over 12 weeks	12

Note. CT = Cognitive Therapy; ERP = Exposure and Response therapy; WLC = Waitlist Control; MI = Motivational Interviewing; N/R = Not Reported; SD = standard decision; TM = thought mapping; Y-BOCS = Yale Brown Obsessive-Compulsive Scale.

[a] based on IPD provided.
[b] Psychoeducation.
[c] Relaxation Control.
[d] Anxiety management training.
[e] Eye Movement Desensitization and Reprocessing.
[f] Pill Placebo + Autogenic Training.
[g] Stress Management Training.

Table 3
Risk of bias of included trials.

Author(s)	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other sources of bias
Anderson and Rees (2007)	Unclear	Unclear	High	Unclear	Low	High	Low
Anderson et al. (2012)	Low	Unclear	High	Low	Low	Low	Low
Baruah et al. (2018)	Low	Low	High	Low	Low	Low	Low
Belloch et al. (2011)	Unclear	Unclear	High	High	High	High	Low
Cordioli et al. (2003)	Low	Low	High	Low	Low	Low	Low
Cottraux et al. (2001)	Unclear	Unclear	High	Unclear	High	Low	High
Gomes et al. (2016)	Low	Unclear	High	Unclear	Low	Low	Low
Hauschildt et al. (2016)	Unclear	Low	High	Low	Low	Low	Low
Herbst et al. (2014)	Low	Unclear	High	Unclear	Low	Low	Low
Jaurietta et al. (2008)	Low	Low	High	Unclear	Low	High	Low
Kobak et al. (2015)	Low	Unclear	High	Low	Unclear	High	High
Lindsay et al. (1997)	Unclear	Unclear	High	Unclear	Unclear	Unclear	Low
Lovell et al. (2017)	Low	Low	High	Low	Low	Low	Low
Lovell et al. (2006)	Low	Low	High	Low	Low	Low	Low
Marsden et al. (2017)	Low	Low	High	Low	Low	Low	Low
McLean (2001)	Low	Unclear	High	Unclear	High	Low	Low
Meyer et al. (2010)	Low	Unclear	High	Low	Low	Low	Low
Nakatani et al. (2005)	Unclear	Low	High	Low	Low	Unclear	Low
Tolin et al. (2007)	Low	Unclear	High	Low	Low	Low	Low
van Oppen et al. (1995)	Unclear	Unclear	High	Unclear	High	High	Low
Vogel et al. (2014)	Unclear	Unclear	High	Low	Unclear	Unclear	Unclear
Vogel et al. (2004)	High	Low	High	High	Low	Unclear	Low
Whittal et al. (2005)	Unclear	Unclear	High	Low	High	Low	Low
Whittal et al. (2010)	Unclear	Unclear	High	Unclear	High	Low	Low

Table 4
Weighted mean dropout rates for each treatment and control condition at post-treatment (≤4 weeks after treatment) and follow-up (3–6 months after treatment).

	Post-treatment		Follow-up	
	k	Missing data% [95% CI]	*k*	Missing data% [95% CI]
Psychological treatment				
Individual CT + ERP	3	28% [17–40%]		
Group CT	2	23% [3–54%]		
Individual ERP	10	15% [8–24%]	7	24% [11–39%]
Group ERP	2	15% [0–44%]	2	15% [0–47%]
Self-help CT + ERP	4	14% [3–30%]	3	27% [9–51%]
Group CT + ERP	6	13% [4–27%]	3	5% [2–10%]
Individual CT	5	12% [8–18%]	2	16% [6–25%]
Self-help ERP	6	2% [0–7%]		
Self-help CT	1			
'Other' treatment	1			
Controls				
WLC	7	9% [1–22%]		
Placebo Control	6	7% [1–15%]	4	16% [11–21%]

Note. *k* = number of conditions; N/R = Not reported; CI = confidence interval; CT = cognitive therapy; ERP = exposure and response prevention; WLC = waitlist control.

3.5. Preliminary analysis

The pre to post within-group effect sizes of RCTs that provided IPD and those that did not provide IPD were not significantly different (Q = 0.02, df = 2, p = .9).

3.6. Clinical significance analysis

3.6.1. Jacobson clinical significance criteria

Recovery. Weighted recovery rates for each treatment and control category are in Table 5. At post-treatment, recovery rates were significantly higher for treated patients (32%) compared to controls (3%). Moderate heterogeneity was indicated for treatment (I^2 = 78%). The highest recovery rates were found for individual CT (55%), followed by group CT + ERP (42%) and individual ERP (41%). Recovery rates were comparable between group CT, individual CT + ERP and self-help ERP (27–28%). The lowest recovery rates were group ERP and self-help CT + ERP (12–16%). Treatment category moderated recovery rates. Inspection of 95% CI indicated that individual CT had significantly higher recovery rates (55%) than two categories; individual CT + ERP (27%) and self-help CT + ERP (16%). Only one other significant difference emerged; recovery rates were significantly higher for individual ERP (41%) than self-help CT + ERP (16%). Recovery rates were not moderated by control category (Table 5).

At follow-up, recovery rates were significantly higher for treated patients (37%) compared to controls (20%) but this difference was not significant. Moderate heterogeneity was indicated for treatment (I^2 = 82%) and controls (I^2 = 79%). Recovery rates were highest for group CT + ERP (61%), closely followed by individual CT (58%). The next highest recovery rates were individual ERP (45%), then group ERP (37%). Self-help CT + ERP had the lowest recovery rate (22%; Table 5). However, treatment category did not moderate recovery rates.

Improvement and deterioration. Weighted improvement and deterioration rates for each treatment and control category are presented in Table 6. At post-treatment, improvement rates were significantly higher for treated patients (50%) compared to controls (6%). Moderate heterogeneity was indicated for treatment (I^2 = 84%) and controls (I^2 = 54%). Improvement rates were highest for group CT + ERP (70%). This was closely followed by group individual CT (68%) and individual ERP (62%). Improvement rates were similar between individual CT + ERP (49%) and group CT (42%), whilst the lowest improvement rates were found for self-help ERP, self-help CT + ERP, group ERP (24–32%). Treatment category moderated improvement rates. Inspection of 95% CIs revealed that self-help CT + ERP (24%) and self-help ERP (32%) had significantly lower improvement rates than individual CT (68%) and individual ERP (62%). Self-help CT + ERP also had significantly lower improvement rates than group CT + ERP (70%). Improvement rates were not moderated by the type of control condition (Table 6).

At follow-up, improvement rates were significantly higher for treated patients (50%) compared to controls (26%) but this difference was not significant. Moderate heterogeneity was indicated for treatment (I^2 = 81%) and controls (I^2 = 78%). Improvement rates were

Table 5
Weighted recovery rates according to Jacobson's clinical significance criteria at post-treatment (≤4 weeks after treatment) and follow-up (3–6 months after treatment).

	Post-treatment						Follow-up					
	k	*n*	Recovery% [95% CI]	I^2	Q(df)	*p*	*k*	*n*	Recovery% [95% CI]	I^2	Q(df)	*p*
Psychological treatment	**40**	**1192**	**34% [27–41%]**	**0.78**			**22**	**716**	**37% [28–47%]**	**0.82**		
Treatment type					23.2(7)	<0.01*					8.96(4)	.06
Individual CT	5	141	55% [45–66%]	0.35			2	63	58% [30–82%]	0.80		
Group CT + ERP	6	197	42% [24–62%]	0.83			3	111	61% [36–81%]	0.83		
Individual ERP	10	220	41% [35–48%]	0.00			7	141	43% [35–52%]	0.00		
Group CT	2	53	28% [6–72%]	0.87								
Individual CT + ERP	3	41	27% [16–43%]	0.00								
Self-help ERP	6	124	27% [18–48%]	0.20								
Self-help CT + ERP	4	290	16% [6–35%]	0.88			3	232	22% [11–39%]	0.85		
Group ERP	2	61	12% [0–83%]	0.84			2	60	37% [26–50%]	0.00		
Self-help CT	1											
'Other' treatment	1				0.13(1)	0.72						
Controls	**13**	**429**	**5% [2–9%]**	0.46			**4**	**154**	**20% [9–39%]**	**0.79**		
Control type												
Placebo Control	6	182	4% [1–18%]	0.66			4	154	20% [9–39%]	0.79		
WLC	7	247	6% [3–10%]	0.00								
Treatment vs. control					30.43(1)	<0.01*					2.66(1)	.10

Note. CI = confidence interval; CT = cognitive therapy; df = degrees of freedom; ERP = exposure and response prevention; *k* = number of conditions; *n* = number of patients; N/R = not reported; WLC = waitlist control.

highest for group CT + ERP (88%), markedly higher than the improvement rates for the next highest categories; individual CT (65%) and individual ERP (51%). The lowest improvement rates were group ERP and self-help CT + ERP (34–40%). Treatment category moderated improvement rates. Inspection of 95% CIs revealed that group CT + ERP (88%) had significantly higher improvement rates than individual ERP (51%), group ERP (40%) and self-help CT + ERP (34%; Table 6).

At post-treatment and follow-up, deterioration rates were identical and did not significantly differ between treated patients (1–3%) and controls (2–3%). Neither treatment category nor control category moderated deterioration rates (Table 6).

3.6.2. Asymptomatic criterion

Asymptomatic. Weighted asymptomatic rates for each treatment and control category are presented in Table 7. Moderate heterogeneity was indicated for treatment (I^2 = 80%) and controls (I^2 = 60%). At post-treatment, asymptomatic rates were significantly higher for treated patients (18%) than controls (4%). The highest asymptomatic rates were found for individual CT (41%), markedly higher than the next highest asymptomatic rates; group CT (25%), group CT + ERP (24%) and individual ERP (20%). Asymptomatic rates were lowest for self-help ERP, self-help CT + ERP, individual CT + ERP, and group ERP (9–12%). Treatment category moderated asymptomatic rates. Inspection of 95% CIs revealed only one significant difference between treatment categories; asymptomatic rates were significantly higher for individual CT (41%) than self-help ERP (11%). The control category did not moderate asymptomatic rates (Table 7).

At follow-up, asymptomatic rates were very similar between treated patients (22%) and controls (20%) and did not significantly differ. Asymptomatic rates remained highest for individual CT (46%), closely followed by group CT + ERP (43%). The next highest asymptomatic rates were found for individual ERP (26%); while group ERP and self-help CT + ERP had the lowest asymptomatic rates (8–11%). Treatment category moderated asymptomatic rates. Inspection of 95% CIs revealed only two significant differences; asymptomatic rates were significantly higher for individual CT (46%) and group CT + ERP than self-help CT + ERP (8%; Table 7).

3.7. Effect size analysis

SMD within-group effects-sizes. Weighted SMD within-group effect sizes for each treatment and control category are presented in Table 8. At post-treatment, effect sizes were significantly larger for treated patients (g = 1.28) than controls (g = 0.3). Moderate heterogeneity was indicated for treatment (I^2 = 91%) and controls (I^2 = 80%). Treatment category moderated effect sizes. Individual CT had the largest effect size (g = 1.85), closely followed by group CT + ERP (g = 1.80), and then individual ERP (g = 1.58). Group ERP, self-help ERP, individual CT + ERP, and self-help CT + ERP all had similar effect sizes (g = 0.9–1.02); while group CT had the smallest effect size (g = 0.41).

At follow-up, effect sizes were significantly larger for treated patients (g = 1.46) than controls (g = 0.91). Moderate heterogeneity was indicated for treatment (I^2 = 87%) and controls (I^2 = 83%). Treatment category moderated effect sizes. Group CT + ERP had the largest effect size (g = 2.90). The next largest effect sizes were individual CT (g = 1.96), individual ERP (g = 1.39), and group ERP (g = 1.27). Self-help CT + ERP had the smallest effect size (g = 1.06; Table 8).

4. Discussion

An IPD-MA of RCTs was conducted to evaluate the efficacy of psychological treatments for OCD. Using Jacobson's clinical significance methodology enabled investigation of recovery, improvement and deterioration rates. The proportion of patients who were almost absent of OCD symptomology (i.e. similar to the levels experienced by individuals without OCD) was determined by using asymptomatic criterion. Results showed that treated patients were more likely than controls to recover and improve at post-treatment but no more likely to improve at follow-up. Treated patients were also more likely to be asymptomatic than controls at post-treatment. However, treated patients were no more likely to be asymptomatic at follow-up. The likelihood of deteriorating did not differ between treated and control patients at post-treatment or follow-up.

To detect whether different methods of analysis suggested different conclusions, we also evaluated treatment efficacy using within-group effect size analysis. In line with findings of previous meta-analyses, large within-group effect sizes were found for treated patients at post-treatment and follow-up, and within-group effect sizes were significantly larger for treated patients than controls. Thus, the different analysis methods largely converged indicating that psychological treatments are more efficacious than controls for OCD. However, the clinical significance analysis added a practical indication of the actual benefit of psychological treatment. Only 32% of treated patients recovered post-treatment compared with 3% of controls, rising to 38%

Table 6
Weighted improvement and deterioration rates according to Jacobson's second criterion (RCI) at post-treatment (≤ 4 weeks after treatment) and follow-up (3–6 months after treatment).

						Post-treatment							Follow-up				
	k	*n*	Improvement% [95% CI]	I^2	Q(df)	*p*	Deterioration% [95% CI]	I^2	Q(df)	*p*	*k*	*n*	Improvement% [95% CI]	I^2	Q(df)	*p*	Deterioration% [95% CI]	I^2	Q(df)	*p*
Psychological treatment	**40**	**1192**	**50% [42–59%]**	**0.84**			**2% [1–3%]**	0.00			**22**	**716**	**50% [40–60%]**	**0.81**			**2% [1–3%]**	**0.00**		
Treatment type					40.8(7)	< 0.01*			3.5(7)	.89					16.4(4)	< 0.01*			0.82(4)	.94
Individual CT	5	141	68% [56–78%]	0.42			2% [1–6%]	0.00			2	63	65% [34–88%]	0.82			2% [0–11%]	0.00		
Group CT + ERP	6	197	70% [46–87%]	0.86			2% [1–6%]	0.00			3	111	88% [63–97%]	0.80			1% [0–7%]	0.00		
Individual ERP	10	220	62% [55–68%]	0.00			2% [0–6%]	0.00			7	141	51% [43–59%]	0.00			3% [1–7%]	0.00		
Individual CT + ERP	3	41	49% [34–64%]	0.00			3% [1–15%]	0.00												
Group CT	2	53	42% [8–86%]	0.91			2% [0–12%]	0.00												
Self-help ERP	6	124	32% [18–50%]	0.61			3% [1–9%]	0.00												
Group ERP	2	61	23% [2–76%]	0.89			2% [0–10%]	0.00			2	60	40% [28–53%]	0.00			2% [0–11%]	0.00		
Self-help CT + ERP	4	290	23% [14–35%]	0.72			1% [0–3%]	0.00			3	232	34% [25–46%]	0.65			1% [0–4%]	0.00		
Self-help CT	1																			
'Other' treatment	1																			
Controls	**13**	**429**	**6% [3–12%]**	**0.54**			**2% [1–5%]**	0.00			**4**	**154**	**26% [13–44%]**	**0.78**			**3% [1–9%]**	**0.00**		
Control type					0.69(1)	0.41			0.06(1)	.81										
Placebo Control	6	182	5% [1–21%]	0.73			2% [1–6%]	0.00			4	154	26% [13–44%]	0.78			3% [1–9%]	0.00		
WLC	7	247	10% [7–15%]	0.00			2% [1–6%]	0.00												
Treatment vs control					45.39(1)	< 0.01*			0.00(1)	.95					4.88(1)	.03*			0.58(1)	.49

Note. CI = confidence interval; CT = cognitive therapy; df = degrees of freedom; ERP = exposure and response prevention; *k* = number of conditions; *n* = number of patients; N/*R* = not reported; tx = treatment; WLC = waitlist control.

and 21% respectively after 3 to 6 months. Furthermore, only 18% of treated patients were asymptomatic at post-treatment compared to 4% of controls; and at follow-up only around 20% of patients were asymptomatic regardless of whether they were allocated to treatment or control conditions. Deterioration rates were very low (0–1%) and did not differ between treated and control patients at either post-treatment or at follow-up.

Unlike previous meta-analyses (Abramowitz et al., 2002; Eddy et al., 2004; Gava et al., 2007; Jónsson and Hougaard, 2009; Olatunji et al., 2013; Öst et al., 2015; Pearcy et al., 2016; Romanelli et al., 2014; Rosa-Alcázar et al., 2008; Schwartze et al., 2016; Skapinakis et al., 2016; Wootton, 2016), our IPD-MA showed that treatment efficacy was moderated by treatment type. In general, the most efficacious treatment was individual CT, followed by group CT + ERP and, to a lesser extent, individual ERP. Amongst these treatments, recovery rates of 55%, 42% and 41% respectively were found at post-treatment; while asymptomatic rates of 41%, 24% and 20% respectively were found at post-treatment.

Relative to patients receiving face-to-face interventions, patients receiving self-help interventions were less likely to improve, recover and become 'asymptomatic' at post-treatment and follow-up. Specifically, recovery rates for self-help interventions at follow-up ranged from 16% to 27%; 'asymptomatic' rates were lower, ranging from 9 to 11%. These results are surprising given that previous meta-analyses conclude that self-help interventions are efficacious (Pearcy et al., 2016; Wootton, 2016) and that face-to-face and remote forms of CBT for OCD are equally efficacious (Dettore et al., 2015).

4.1. Why were treatments largely ineffective?

Collectively, data indicate that most treated patients benefit from psychological therapy, with those receiving individual CT making the greatest gains. However, in practical terms, our findings indicate that even the most effective treatments leave 60% of patients experiencing potentially distressing OCD symptomatology. More worryingly, some self-help interventions leave up to 90% of patients distressed 3–6 months post-intervention. Thus, there remains considerable room for improvement with regards to treatment efficacy.

The low recovery and asymptomatic rates found in the current study raise three important questions. First, why were recovery and asymptomatic rates so low? Second, what is acceptable in order for an intervention to be classed as effective; in other words, should we accept current recovery rates given that a score of 12 on the Y-BOCS can indicate significant OCD symptoms (Fisher and Wells, 2005)? Third, what are the implications for future practice and research?

It is difficult to address the first question without speculation. We know little about treatment fidelity or specific patient or intervention characteristics beyond those reported in the included studies, nor how these may have influenced treatment delivery or outcome. For example, we do not know if different types of obsessions impacted on the efficacy of group interventions. Furthermore, it is plausible that self-help interventions may not have been intensive enough for the samples studied; the mean Y-BOCS scores of patients allocated to self-help interventions ranged from 18 to 31, indicative of moderate to severe OCD, yet NICE's stepped care model (National Institute for Health and Care Excellence [NICE], 2005) recommends brief telephone or self-help CT + ERP for individuals with mild OCD symptoms only. This may account for why lower recovery and 'asymptomatic' rates were found for self-help interventions in this IPD-MA. Finally, it is possible that different results would have been found if studies evaluating the efficacy of non-manualised psychological interventions had been included. Regardless, data indicate that recommended psychological approaches increased recovery over control conditions by only 17%, and had no impact on 'asymptomatic' rates - findings very similar to those observed by Fisher and Wells (2005) 15 years ago.

With regards question two, this is a value judgement but it seems

Table 7
Weighted asymptomatic rates according to according to the asymptomatic criterion (YBOCS ≤ 7) at post-treatment (≤ 4 weeks after treatment) and follow-up (3–6 months after treatment).

	Post-treatment						Follow-up					
	k	*n*	Asymptomatic% [95% CI]	I^2	Q(df)	*p*	*k*	*n*	Asymptomatic% [95% CI]	I^2	Q(df)	*p*
Psychological treatment	**40**	**1192**	**18% [14–24%]**	0.80			**22**	716	**22% [15–30%]**	**0.77**		
Treatment types					15.1(7)	.04*					12.54(4)	< 0.01*
Individual CT	5	141	41% [21–64%]				2	63	46% [32–61%]	0.29		
Group CT	2	53	25% [7–60%]	0.80								
Group CT + ERP	6	197	24% [16–34%]	0.47			3	111	43% [32–55%]	0.33		
Individual ERP	10	220	20% [15–26%]	0.00			7	141	26% [16–38%]	0.44		
Individual CT + ERP	3	41	11% [4–25%]	0.00								
Self-help CT + ERP	4	290	9% [3–26%]	0.85			3	232	8% [2–28%]	0.88		
Self-help ERP	6	124	11% [7–19%]	0.00								
Group ERP	2	61	9% [1–58%]	0.72			2	60	11% [0–79%]	0.82		
Self-help CT	1											
Other	1											
Controls	**13**	**429**	**4% [1–12%]**	0.60			**4**	**154**	**20% [6–49%]**	**0.88**		
Control type					0.78(1)	.38						
Placebo Control	6	182	6% [1–32%]	0.85			4	154	20% [6–49%]	0.88		
WLC	7	247	2% [1–6%]	0.00								
Treatment vs. control					6.84(1)	< 0.01*					0.03(1)	.88

Note. CI = confidence interval; CT = cognitive therapy; df = degrees of freedom; ERP = exposure and response prevention; *k* = number of conditions; *n* = number of patients; N/*R* = not reported; WLC = waitlist control.

reasonable to conclude that, when only around one-fifth of patients are asymptomatic following psychological treatment in general for OCD, researchers and clinicians should challenge the assertion that highly effective psychological interventions for OCD exist (particularly given that we may even have overestimated the efficacy of psychological treatments because we conducted analyses on a per-protocol rather than ITT basis (Cuijpers et al., 2010)).

When considering question three, we argue that, prior to developing more cost-effective means of delivering existing treatments, there is a need to develop more efficacious interventions for people with OCD. Currently, NICE guidance (2005) advocates a stepped-care approach to treating OCD, in which access to psychological and pharmacological interventions is determined based, in part, on the severity of patients' symptoms. Whilst this review was able to examine moderator effects of treatment type, there is obvious clinical utility in establishing patient characteristics that are associated with differential treatment responses. However, this is an area that has been largely overlooked in previous meta-analyses. Over 50 years ago, Paul (1967) argued that: "the question towards which all outcome research should effectively be directed is the following: What treatment, by whom, is most effective for this individual with that specific problem, and under which set of circumstances?' (p.111). Paul's question represents the first challenge to researchers, given the changing landscape of healthcare and increasing financial pressures upon healthcare providers, and the lack of progress since Fisher and Wells' original clinical significance review (Fisher and Wells, 2005). We also advocate that researchers go 'back to basics', and focus on model-specific research designed to identify and test psychological mechanisms that maintain distress, together with patient moderators of treatment response (Steketee et al., 2019). More recently, major advances in cancer treatment have been achieved following the introduction of 'precision medicine' (Schwaederle et al., 2015) – an approach designed to match individuals to individual treatments based on their characteristics (Hamburg and Collins, 2010). Adoption of a similar approach may help to improve the overall efficacy of interventions for OCD, and may shift focus towards more idiosyncratic psychological interventions for OCD.

4.2. Strengths and limitations

This is the first study to use a standardised recovery definition to examine the extent to which recovery and asymptomatic rates for psychological therapies for OCD 'map onto' effect size estimates.

Table 8
Within-group effect sizes at post-treatment (≤ 4 weeks after treatment) and follow-up (3–6 months after treatment).

	Post-treatment							Follow-up						
	k	*n*	Hedge's g [95% CI]	*p*	I^2	Q(df)	*p*	k	N	Hedge's g [95% CI]	*p*	I^2	Q(df)	*p*
Psychological treatment	**40**	**1194**	**1.28 [1.07, 1.49]**	**< 0.01**	**0.91**			**22**	**732**	**1.46 [1.18, 1.74]**	**< 0.01***	**0.87**		
Treatment type						34.15 (7)	< 0.01*						14.57 (6)	< 0.02*
Individual CT	5	141	1.85 [1.45, 2.26]	< 0.01*	0.43			2	63	1.96 [0.75, 3.18]	< 0.01*	0.81		
Group CT + ERP	6	197	1.80 [1.08, 2.53]	< 0.01*	0.92			3	111	2.90 [1.64, 4.15]	< 0.01*	0.86		
Individual ERP	10	222	1.58 [1.38, 1.77]	< 0.01*	0			7	141	1.39 [1.13, 1.66]	< 0.01*	0.26		
Group ERP	2	61	1.02 [0.46, 1.58]	< 0.01*	0.73			2	60	1.27 [0.95, 1.60]	< 0.01*	0		
Self-help ERP	6	124	1.02 [0.6, 1.44]	< 0.01*	0.72									
Self-help CT + ERP	4	290	0.91 [0.58, 1.24]	< 0.01*	0.83			3	232	1.06 [0.76, 1.36]	< 0.01*	0.70		
Individual CT + ERP	3	41	0.9 [0.58, 1.22]	< 0.01*	0.49									
Group CT	2	53	0.41 [−0.50,1.33]	.38	0.97									
Self-help CT	1													
'Other' treatment	1													
Control	**13**	**429**	**0.30 [0.14, 0.46]**	**< 0.01***	**0.8**			**4**	**154**	**0.91 [0.46, 1.36]**	**< 0.01***	**0.83**		
Placebo Control	6	182	0.44 [0.12, 0.75]	< 0.01*	0.88			4	154	0.90 [0.46, 1.35]	< 0.01*	0.89		
WLC	7	247	0.20 [0.04, 0.20]	0.01*	0.63									
Treatment vs. control						53.43 (1)	< 0.01*						4.17(1)	.04*

Note. CI = confidence interval; CT = cognitive therapy; df = degrees of freedom; ERP = exposure and response prevention; *k* = number of conditions; *n* = number of patients; N/*R* = not reported; WLC = waitlist control.

However, our study has limitations. First, not all authors provided IPD, therefore data from only 24 of the 43 eligible RCTs (representing IPD for 1626 of 2455 patients) were analysed. However, there was no difference in pre-post within-group effect sizes between the RCTs that provided data and those that did not, suggesting that the included RCTs were representative of published trials using the Y-BOCS as an outcome measure. Second, only RCTs that used either the self-report or clinician-rated Y-BOCS – the 'gold standard' and most widely used measure of OCD severity (Kyrios et al., 2015) - as an outcome measure were included. This means conclusions about the effectiveness of psychological interventions evaluated using other measures of OCD symptom severity cannot be drawn, nor can conclusions regarding the wider effectiveness of psychological interventions for adults with OCD (e.g. impact on anxiety or depressive symptomatology or suicidal ideation, or on broader indices of functioning such as quality of life or family or job role functioning). A more fine-grained analysis of factors associated with OCD severity is needed in order to better understand the wide-ranging impact of OCD. Finally, we were unable to consider the long-term efficacy of treatments because only two RCTs provided long-term (>12 month) follow-up data. Furthermore, small sample sizes and lack of follow-up data for some interventions potentially limits the conclusions that can be drawn about their long-term efficacy.

5. Conclusions

Current psychological interventions for OCD do not lead to deterioration of symptoms, but do not completely alleviate distressing symptoms. Around 80% of patients with OCD remain distressed following manualised psychological treatment, and only 60% can be classed as recovered. Although individual CT shows promise, more efficacious psychological treatments are needed for patients with OCD. Rather than continue to evaluate different modes of delivery of existing treatments, researchers should focus on testing alternative theoretical models, together with patient moderators of treatment response.

CRediT authorship contribution statement

P.L. Fisher: Conceptualization, Data curation, Methodology, Formal analysis, Supervision, Writing - review & editing. **M.G. Cherry:** Formal analysis, Methodology, Supervision, Writing - original draft, Writing - review & editing. **T. Stuart:** Data curation, Formal analysis. **J.W. Rigby:** Formal analysis, Writing - review & editing. **J. Temple:** Methodology, Formal analysis, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

None.

Acknowledgements

We thank the authors of included studies for providing individual patient data for analysis.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jad.2020.06.019.

References

Abramowitz, J.S., Franklin, M.E., Foa, E.B., 2002. Empirical status of cognitive-behavioral therapy for obsessive-compulsive disorder: a meta-analytic review. Rom. J. Cogn. Behav. Psychother. 89–104. https://doi.org/10.1002/14651858.CD005333.pub2.
American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). American Psychiatric Pub.
American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders, (4th Text Revision ed.). Washington, DC.
American Psychiatric Association, Koran, L. M., Hanna, G. L., Hollander, E., Nestadt, G., and Simpson, H. B. (2007). Practice guideline for the treatment of patients with obsessive-compulsive disorder.
Baer, L., Minichiello, W.E., 2008. Reasons for inadequate utilization of cognitive-behavioral therapy for obsessive-compulsive disorder. J. Clin. Psychiatry 69 (4), 676.
Barlow, D.H., 2010. Negative effects from psychological treatments: a perspective. Am. Psychol. 65 (1), 13.
Bobes, J., Gonzalez, M., Bascaran, M., Arango, C., Saiz, P., Bousono, M., 2001. Quality of life and disability in patients with obsessive-compulsive disorder. Eur. Psychiatry 16 (4), 239–245.
Borenstein, M., Hedges, L., Higgins, J., Rothstein, H., 2009. Introduction to Meta-Analysis. John Wiley & Sons, Ltd., Chichester.
Bravata, D.M., Olkin, I., 2001. Simple pooling versus combining in meta-analysis. Eval. Health Prof. 24 (2), 218–230.
Centre for Reviews and Dissemination. (2009). *Systematic Reviews: CRD's Guidance For Undertaking Reviews in Healthcare*. Retrieved fromhttp://www.york.ac.uk/crd/SysRev/!SSL!/WebHelp/SysRev3.htm.
Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. Routledge Academic, New York, NY.
Cuijpers, P., van Straten, A., Bohlmeijer, E., Hollon, S.D., Andersson, G., 2010. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. In: Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews [Internet]. Centre for Reviews and Dissemination (UK).
Cottraux, J., Note, I., Yao, S.N., Lafont, S., Note, B., Mollard, E., ... Dartigues, J.-.F., 2001. A randomized controlled trial of cognitive therapy versus intensive behavior therapy in obsessive compulsive disorder. Psychother. Psychosom. 70 (6), 288–297.
Cuijpers, P., Reijnders, M., Karyotaki, E., de Wit, L., Ebert, D.D., 2018. Negative effects of psychotherapies for adult depression: a meta-analysis of deterioration rates. J. Affect. Disord. 239, 138–145.
de la Cruz, L.F., Rydell, M., Runeson, B., D'Onofrio, B.M., Brander, G., Rück, C., ... Mataix-Cols, D., 2017. Suicide in obsessive–compulsive disorder: a population-based study of 36 788 Swedish patients. Mol. Psychiatry 22 (11), 1626.
De Putter, L.M., Koster, E.H., 2017. The effects of obsessive-compulsive symptoms and disorder-relevant stimuli on the dynamics of selective attention. J. Obsessive Compuls. Relat. Disord. 15, 74–84.
DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. Control Clin. Trials 7 (3), 177–188.
Dettore, D., Pozza, A., Andersson, G., 2015. Efficacy of technology-delivered cognitive behavioural therapy for OCD versus control conditions, and in comparison with therapist-administered CBT: meta-analysis of randomized controlled trials. Cogn. Behav. Ther. 44 (3), 190–211.
Eddy, K.T., Dutra, L., Bradley, R., Westen, D., 2004. A multidimensional meta-analysis of psychotherapy and pharmacotherapy for obsessive-compulsive disorder. Clin. Psychol. Rev. 24 (8), 1011–1030.
Eisen, J.L., Mancebo, M.A., Pinto, A., Coles, M.E., Pagano, M.E., Stout, R., Rasmussen, S.A., 2006. Impact of obsessive-compulsive disorder on quality of life. Compr. Psychiatry 47 (4), 270–275.
Federici, A., Summerfeldt, L.J., Harrington, J.L., McCabe, R.E., Purdon, C.L., Rowa, K., Antony, M.M., 2010. Consistency between self-report and clinician-administered versions of the Yale-Brown obsessive–compulsive scale. J. Anxiety Disord. 24 (7), 729–733.
Fisher, P., Wells, A., 2005. How effective are cognitive and behavioral treatments for obsessive–compulsive disorder? A clinical significance analysis. Behav. Res. Ther. 43 (12), 1543–1558.
Foa, E., Liebowitz, M., Kozak, M., Davies, S., Campeas, R., Franklin, M., ... Tu, X., 2005. Randomized, placebo-controlled trial of exposure and ritual prevention, clomipramine, and their combination in the treatment of obsessive-compulsive disorder. Am. J. Psychiatry 162 (1), 151–161.
Gava, I., Barbui, C., Aguglia, E., Carlino, D., Churchill, R., De Vanna, M., McGuire, H., 2007. Psychological treatments versus treatment as usual for obsessive compulsive disorder (OCD). Cochrane Database Syst. Rev.(2).
Goodman, W.K., Price, L.H., Rasmussen, S.A., Mazure, C., Delgado, P., Heninger, G.R., Charney, D.S., 1989a. The Yale-Brown obsessive compulsive scale: II. Validity. Arch. Gen. Psychiatry 46 (11), 1012–1016.
Goodman, W.K., Price, L.H., Rasmussen, S.A., Mazure, C., Fleischmann, R.L., Hill, C.L., ... Charney, D.S., 1989b. The Yale-Brown obsessive compulsive scale: I. Development, use, and reliability. Arch. Gen. Psychiatry 46 (11), 1006–1011.
Goodwin, R., Koenen, K.C., Hellman, F., Guardino, M., Struening, E., 2002. Helpseeking and access to mental health treatment for obsessive-compulsive disorder. Acta Psychiatr. Scand. 106 (2), 143–149.
Grabill, K., Merlo, L., Duke, D., Harford, K.-.L., Keeley, M.L., Geffken, G.R., Storch, E.A., 2008. Assessment of obsessive–compulsive disorder: a review. J. Anxiety Disord. 22 (1), 1–17.
Hamburg, M., Collins, F., 2010. The path to personalized medicine. N. Engl. J. Med. 363, 301–304.
Hedges, L.V., Shymansky, J.A., Woodworth, G., 1989. A Practical Guide to Modern

Methods of Meta-Analysis. National Science Teachers Association, Washington, D.C.
Higgins, J.P.T., & Green, S. (2011). *Cochrane Handbook For Systematic Reviews of Interventions, version 5.1.0*. Retrieved fromwww.handbook.cochrane.org.
Higgins, J.P., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. Bmj 327 (7414), 557–560.
Hollander, E., Greenwald, S., Neville, D., Johnson, J., Hornig, C.D., Weissman, M.M., 1998. Uncomplicated and comorbid obsessive-compulsive disorder in an epidemiologic sample. CNS Spectr. 3 (S1), 10–18.
Jacobson, N.S., Follette, W.C., Revenstorf, D., 1984. Psychotherapy outcome research: methods for reporting variability and evaluating clinical significance. Behav. Ther. 15 (4), 336–352.
Jacobson, N.S., Truax, P., 1991. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J. Consult. Clin. Psychol. 59 (1), 12.
Jones, M.K., Wootton, B.M., Vaccaro, L.D., 2012. The efficacy of danger ideation reduction therapy for an 86-year old man with a 63-year history of obsessive-compulsive disorder: a case study. Int. J. Psychol. Behav. Sci. 1, 1–7.
Jónsson, H., Hougaard, E., 2009. Group cognitive behavioural therapy for obsessive–compulsive disorder: a systematic review and meta-analysis. Acta Psychiatr. Scand. 119 (2), 98–106.
Katzman, M.A., Bleau, P., Blier, P., Chokka, P., Kjernisted, K., Van Ameringen, M., 2014. Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. BMC Psychiatry 14 (1), S1.
Keller, M.B., 2003. Past, present, and future directions for defining optimal treatment outcome in depression: remission and beyond. JAMA 289 (23), 3152–3160.
Kyrios, M., Hordern, C., Fassnacht, D.B., 2015. Predictors of response to cognitive behaviour therapy for obsessive-compulsive disorder. Int. J. Clin. Health Psychol. 15 (3), 181–190.
Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., ... Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 6 (7), e1000100.
Lunnen, K.M., Ogles, B.M., 1998. A multiperspective, multivariable evaluation of reliable change. J. Consult. Clin. Psychol. 66 (2), 400.
Marques, L., LeBlanc, N.J., Weingarden, H.M., Timpano, K.R., Jenike, M., Wilhelm, S., 2010. Barriers to treatment and service utilization in an internet sample of individuals with obsessive–compulsive symptoms. Depress. Anxiety 27 (5), 470–475.
National Institute for Health and Care Excellence. (2005). *Obsessive-Compulsive Disorder and Body Dysmorphic Disorder: Treatment*. Retrieved from London, UK: https://www.nice.org.uk/guidance/CG31/chapter/1-Guidance#steps-35-treatment-options-for-people-with-ocd-or-bdd.
Ogles, B.M., Lunnen, K.M., Bonesteel, K., 2001. Clinical significance: history, application, and current practice. Clin. Psychol. Rev. 21 (3), 421–446.
Olatunji, B.O., Davis, M.L., Powers, M.B., Smits, J.A., 2013. Cognitive-behavioral therapy for obsessive-compulsive disorder: a meta-analysis of treatment outcome and moderators. J. Psychiatr. Res. 47 (1), 33–41.
Olatunji, B.O., Deacon, B.J., Abramowitz, J.S., 2009. The cruelest cure? Ethical issues in the implementation of exposure-based treatments. Cogn. Behav. Pract. 16 (2), 172–180.
Öst, L.-.G., Havnen, A., Hansen, B., Kvale, G., 2015. Cognitive behavioral treatments of obsessive–compulsive disorder. A systematic review and meta-analysis of studies published 1993–2014. Clin. Psychol. Rev. 40, 156–169.
Pallanti, S., Hollander, E., Bienstock, C., Koran, L., Leckman, J., Marazziti, D., Pato, M., Stein, D., Zohar, J., the International Treatment Refractory OCD Consortium, 2002. Treatment non-response in OCD: methodological issues and operational definitions. Int. J. Neuropsychopharmacol. 5, 181–191.
Paul, G., 1967. Strategy of outcome research in psychotherapy. J. Consult. Psychol. 31 (2), 109–118.
Pearcy, C.P., Anderson, R.A., Egan, S.J., Rees, C.S., 2016. A systematic review and meta-analysis of self-help therapeutic interventions for obsessive–compulsive disorder: is therapeutic contact key to overall improvement? J. Behav. Ther. Exp. Psychiatry 51, 74–83.
Pozza, A., Dettore, D., 2017. Drop-out and efficacy of group versus individual cognitive behavioural therapy: what works best for Obsessive-Compulsive Disorder? A systematic review and meta-analysis of direct comparisons. Psychiatry Res. 258, 24–36.
Riley, R.D., Lambert, P.C., Abo-Zaid, G., 2010. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ 340, c221.
Romanelli, R.J., Wu, F.M., Gamba, R., Mojtabai, R., Segal, J.B., 2014. Behavioral therapy and serotonin reuptake inhibitor pharmacotherapy in the treatment of obsessive–compulsive disorder: a systematic review and meta-analysis of head-to-head randomized controlled trials. Depress. Anxiety 31 (8), 641–652.
Ronk, F.R., Hooke, G.R., Page, A.C., 2016. Validity of clinically significant change classifications yielded by Jacobson-Truax and Hageman-Arrindell methods. BMC Psychiatry 16 (1), 187.
Rosa-Alcázar, A.I., Sánchez-Meca, J., Gómez-Conesa, A., Marín-Martínez, F., 2008. Psychological treatment of obsessive–compulsive disorder: a meta-analysis. Clin. Psychol. Rev. 28 (8), 1310–1325.
Rosenfeld, R., Dar, R., Anderson, D., Kobak, K.A., Greist, J.H., 1992. A computer-administered version of the Yale-Brown obsessive-compulsive scale. Psychol. Assess. 4 (3), 329.
Rozental, A., Castonguay, L., Dimidjian, S., Lambert, M., Shafran, R., Andersson, G., Carlbring, P., 2018. Negative effects in psychotherapy: commentary and recommendations for future research and clinical practice. Br. J. Psychiatry Open 4 (4), 307–312.
Rozental, A., Magnusson, K., Boettcher, J., Andersson, G., 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 (2), 160.
Schwaederle, M., Zhao, M., Lee, J., Eggermont, A., Schilsky, R., Mendelsohn, J., ... Kurzrock, R., 2015. Impact of precision medicine in diverse cancers: a meta-analysis of phase II clinical trials. J. Clin. Oncol. 33 (32), 3817–3825.
Schwartze, D., Barkowski, S., Burlingame, G., Strauss, B., Rosendahl, J., 2016. Efficacy of group psychotherapy for obsessive-compulsive disorder: a meta-analysis of randomized controlled trials. J. Obsessive Compuls. Relat. Disord. 10, 49–61.
Skapinakis, P., Caldwell, D., Hollingworth, W., Bryden, P., Fineberg, N., Salkovskis, P., ... Lewis, G., 2016. Pharmacological and psychotherapeutic interventions for management of obsessive-compulsive disorder in adults: a systematic review and network meta-analysis. Lancet Psychiatry 3 (8), 730–739.
Steketee, G., Frost, R., Bogart, K., 1996. The Yale-Brown obsessive compulsive scale: interview versus self-report. Behav. Res. Ther. 34 (8), 675–684.
Steketee, G., Siev, J., Yovel, I., Lit, K., Wilhem, S., 2019. Predictors and moderators of cognitive and behavioral therapy outcomes for OCD: a patient-level mega-analysis of eight sites. Behav. Ther. 50 (1), 165–176.
Stewart, L.A., Parmar, M.K., 1993. Meta-analysis of the literature or of individual patient data: is there a difference? Lancet 341 (8842), 418–422.
Stewart, L.A., Tierney, J.F., 2002. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. Eval. Health Prof. 25 (1), 76–97.
Stewart, L.A., Clarke, M., Rovers, M., Riley, R.D., Simmonds, M., ... Tierney, J.F., 2015. PRISMA-IPD development group. Preferred reporting items for systematic review and meta-analyses of individual participant data: the PRISMA-IPD Statement. JAMA 313 (16), 1657–1665.
Stroup, D., Berlin, J., Morton, S., Olkin, I., Williamson, G., Rennie, D., ... Thacker, S., 2000. Meta-analysis of observational studies in epidemiology: a proposal for reporting. JAMA 283 (15), 2008–2012.
Temple, J., Salmon, P., Tudur-Smith, C., Huntley, C., Fisher, P., 2018. A systematic review of the quality of randomized controlled trials of psychological treatments for emotional distress in breast cancer. J. Psychosom. Res. 108, 22–31.
Vaccaro, L.D., Jones, M.K., Menzies, R.G., Wootton, B.M., 2010. Danger ideation reduction therapy for obsessive–compulsive checking: preliminary findings. Cogn. Behav. Ther. 39 (4), 293–301.
Vaccaro, L.D., Jones, M.K., Menzies, R.G., Wootton, B.M., 2014. The treatment of obsessive-compulsive checking: a randomised trial comparing danger ideation reduction therapy with exposure and response prevention. Clin. Psychol. 18 (2), 74–95.
Veale, D., Page, N., Woodward, E., Salkovskis, P., 2015. Imagery Rescripting for Obsessive Compulsive Disorder: a single case experimental design in 12 cases. J. Behav. Ther. Exp. Psychiatry 49, 230–236.
Woody, S.R., Steketee, G., Chambless, D.L., 1995. Reliability and validity of the Yale-Brown obsessive-compulsive scale. Behav. Res. Ther. 33 (5), 597–605. https://doi.org/10.1016/0005-7967(94)00076-v.
Wootton, B.M., 2016. Remote cognitive–behavior therapy for obsessive–compulsive symptoms: a meta-analysis. Clin. Psychol. Rev. 43, 103–113.
Ruscio 1, A M, Stein, D J, Chiu, W T, Kessler, R C, 2010 Jan. The Epidemiology of Obsessive-Compulsive Disorder in the National Comorbidity Survey Replication. Mol Psychiatry 15 (1), 53–63. https://doi.org/10.1038/mp.2008.94. Epub 2008 Aug 26.
McHugh, R.K., Whitton, Sarah W, Peckham, Andrew D, Welge, Jeffrey A, Otto, Michael W, 2013 Jun. Patient Preference for Psychological vs Pharmacologic Treatment of Psychiatric Disorders: A Meta-Analytic Review. J Clin Psychiatry 74 (6), 595–602. https://doi.org/10.4088/JCP.12r07757.