Towards personalising treatment

a systematic review and meta-analysis of face-to-face efficacy moderators of cognitive-behavioral therapy and interpersonal psychotherapy for major depressive disorder

Whiston, Aoife; Bockting, Claudi L H; Semkovska, Maria

Published in:
Psychological Medicine

DOI:
10.1017/S0033291719002812

Publication date:
2019

Document version:
Accepted manuscript

Citation for pulished version (APA):
Whiston, A., Bockting, C. L. H., & Semkovska, M. (2019). Towards personalising treatment: a systematic review and meta-analysis of face-to-face efficacy moderators of cognitive-behavioral therapy and interpersonal psychotherapy for major depressive disorder. *Psychological Medicine*, *49*(16), 2657-2668. https://doi.org/10.1017/S0033291719002812

Towards Personalising Treatment: a Systematic Review and Meta-analysis of Face-to-Face Efficacy Moderators of Cognitive Behavioral Therapy and Interpersonal Psychotherapy for Major Depressive Disorder

Aoife Whiston[a] Claudi LH Bockting[b] Maria Semkovska[a,c*]

[a]Department of Psychology, University of Limerick, Limerick, Ireland

[b]Department of Psychiatry, Academic Medical Centre, University of Amsterdam, Netherlands

[c]Health Research Institute, University of Limerick, Limerick, Ireland

* Corresponding author:
Dr Maria Semkovska

Word count: 5080

Running head: Moderators of CBT and IPT efficacy MDD

Abstract

Background. Consistent evidence suggests that face-to-face Cognitive Behavioural Therapy (CBT) and Interpersonal Psychotherapy (IPT) may be equally effective depression treatments. Current clinical research focuses on detecting the best predictors-moderators of efficacy to guide treatment personalisation. However, individual moderator studies show inconsistent findings. This systematic review and meta-analysis aimed to compare the efficacy of CBT and IPT, including combined treatment with antidepressants for depression, and evaluate the predictive power of demographic, clinical presentation and treatment characteristics moderators for both therapies.

Methods. PsycArticles, PsycINFO, PubMed and Cochrane Library were systematically searched through December 2017 for studies that have assessed individuals with major depression receiving either CBT or IPT in a face-to-face format both at pre- and post-treatment. Random-effects moderator meta-analyses were conducted.

Results. One hundred and sixty-eight samples from 137 studies including 11374 participants qualified for the meta-analytic review. CBT and IPT were equally effective across all but one prespecified moderators. For psychotherapy delivered without concomitant antidepressant treatment (ADM's), CBT was superior to IPT (g=1.68, $Q_{between}$ p=0.037). Within-CBT moderator analyses showed that increased CBT efficacy was associated with lower age, high initial depression severity, individual format of administration and no adjunctive ADM's. Within-IPT analyses showed comparable efficacy across all moderators.

Conclusions. Clinical guidance around combined treatment (psychotherapy plus ADM's) should be reconsidered. CBT alone is superior to IPT alone and to combined treatment, while IPT alone is non-inferior to combined treatment. More research is needed to assess the moderating effect of older age and number of previous episodes on IPT efficacy.

Keywords: major depression, cognitive behavioural therapy, interpersonal psychotherapy, face-to-face therapy, efficacy moderator, systematic review, meta-analysis

Introduction

Major depressive disorder (MDD), the most common mental health disorder, is currently the leading cause of disability worldwide with more than 300 million people affected (WHO, 2017). It is associated with high levels of comorbidity (e.g., anxiety and substance use disorders), leaving a pure MDD diagnosis accounting for only one-quarter of all diagnosed patients (Kessler et al., 1996). MDD is also a highly recurrent disorder: those experiencing first episode depression have a 50 to 60% risk of developing a second episode, while relapse estimates reach 70% and 90 % following a second and third episode respectively (Burcusa et al., 2007). To prevent likely relapse, clinicians recommend the use of psychological and/or pharmacological interventions. Uptake of pharmacological interventions has increased dramatically following the advent of third-generation antidepressant medications (ADM's), leading practitioners guidelines to recommend their use as a first-line treatment for severe MDD (NICE, 2009). Despite ADM efficacy, patients tend to prefer psychological interventions alone (McHugh et al., 2013), because of potential side-effects, withdrawal symptoms – often leading to post-withdrawal disorders and high costs – ADM's are 23% more expensive in comparison to psychological interventions (Churchill et al., 2000; Johnstone, 2003; Fava, 2003; Butler et al., 2006; Bockting et al., 2018, Fava et al., 2018) .

Thus, psychological interventions represent key alternative treatments for depression. Among these, systematic reviews (e.g., Cuijpers et al., 2011; Shinohara et al., 2013) and, more recently, network meta-analyses (Barth et al., 2013) have demonstrated that seven therapy types – cognitive-behavioural, non-directive supportive, behavioural activation, psychodynamic, problem-solving, interpersonal psychotherapy, and social skills training – show comparable, moderate to large effects in treating depression. This equal effectiveness of all therapies is commonly referred to as the Dodo Bird verdict. Originating from Lewis Carroll's 'Alice's Adventures in Wonderland', the Dodo bird's announcement of 'everyone has

won, all must have prizes' translates the current status of these psychotherapies (Honyashiki et al., 2014). This effectiveness equivalency finding, along with the high relapse rates consecutive to all depression therapies, has led to a research move towards identifying moderators of treatment response to inform practice and clinical guidance (NICE, 2009). Understanding moderators, that are the clinical and socio-demographic factors influencing therapy efficacy, is essential to optimise personalisation of treatment. Indeed, moderators can be used prescriptively to indicate who is going to respond better to one treatment over another (Fournier et al., 2009). Recent studies increasingly focus on assessment of such moderators of efficacy in otherwise comparable psychological treatments. For example, Driessen and colleagues (2016) conducted post-hoc analyses following a RCT and demonstrated that cognitive-behavioural therapy (CBT) was more beneficial than psychodynamic therapy for a depressive episode that lasted for less than one year, when the inverse pattern was observed for longer episodes: psychodynamic treatment was then more efficacious than CBT. The study also found that, in combination with medication, psychodynamic therapy was more efficacious than CBT for moderately or severely depressed patients. However, to conduct conclusive moderators testing, large sample sizes are needed to achieve the required power (Cuijpers et al., 2016). Among the most widely used psychotherapies are CBT and interpersonal therapy (IPT) (Rucci et al., 2011; Bayliss et al., 2015). CBT, conceptualising MDD as the consequence of maladaptive cognitive processes and related behaviours, has received strong support from both efficacy (Hofmann et al., 2012) and effectiveness research (Butler et al., 2006). IPT, designed specifically for depression treatment, frames the disorder as the consequence of current interpersonal issues and has also a strong empirical support for efficacy (Barth et al., 2013; Law, 2011). Recent meta-analyses suggested that IPT may be more effective than other therapies for certain clinical presentations, further indicating the need for moderators of efficacy research (Cuijpers et al., 2011).

Comparing these therapies in single RCT's has led to mixed findings, with some studies concluding that CBT is more effective than IPT long-term (Shapiro et al., 1994; Rossello et al., 2008), while others finding them of comparable efficacy (Luty et al., 2007; Lemmens et al., 2011). Meta-analytical studies tend to agree that neither is superior in treating depression (Miranda et al., 2005; Cuijpers et al., 2006; Barth et al., 2013). Recent evidence suggests that depending on the MDD presentation, either CBT or IPT can be the most efficacious (Drissen et al., 2016). For example, greater initial depression severity predicts a better response to IPT relative to CBT (Elkin et al., 1995) while comorbid personality disorder and attachment disorder– a better response to CBT (McBride et al., 2006; Carter et al., 2011). Treatment format has also been found a possible moderator, with individual CBT being shown as more effective than group CBT(Cuijpers et al., 2008; Craigie et al., 2009). Some suggest that CBT and IPT are just as effective alone, as they are in adjunctive to ADMs (Thase et al., 1997). Although clinical guidelines recommend the use of combined treatment, no significant trend towards the benefit of additional ADM's is observed in evaluative research (Hollon et al., 2005; Otto et al., 2005; Mintz, 2006). Socio-demographic patients' characteristics also moderate treatment efficacy with female gender and increased age being associated with poorer CBT response (Thase et al., 1994; Hyer et al., 2004) and IPT being suggested as a better alternative to CBT for geriatric depression (Hollon et al., 2005).

To date, only one systematic review sought to identify potential moderators of both CBT and IPT effectiveness (Zhou et al., 2017). It included solely RCT's directly comparing the two treatments. Given the small number of studies included (n=10), only one moderator could be assessed, i.e. study format (individual vs group), but its effect was not significant. Although RCT's are the gold standard for effectiveness assessment, they present a number of limitations for moderators' research. Firstly, the smaller overall population limits the statistical power associated with predictor by treatment interaction effects from ANOVA, multiple and logistic regression models (Fournier et al., 2009). Secondly, as RCT's tend to study depression without

its comorbidities, representative MDD individuals with comorbidities are often excluded to maximise homogeneity, limiting the evaluation of that potential moderator (Budd et al., 2009). Thirdly, restraining reviews to only direct comparisons does not lead to a representative sample of studies of either treatments, leading to questionable conclusions (Gartlehner et al., 2008). Combining studies through a meta-analytical approach allows to optimise power and compare efficacy beyond the limitations of research considering only RCT's of direct comparisons (Borenstein et al., 2009).

Consequently, the present study aimed to: (a) compare the overall efficacy of CBT and IPT for depression through a comprehensive systematic review, not limited to direct comparisons; and (b) evaluate, through a meta-analysis, the effects of commonly computed preselected moderators of therapy efficacy on face-to-face CBT and IPT.

Methods

PRISMA guidelines for conducting and reporting systematic reviews were followed (Moher et al., 2009).

Search Strategy and selection criteria

The electronic databases PsycArticles, PsycINFO, PubMed and Cochrane Library were searched from the year 1980 (MDD diagnosis current conceptualisation [APA, 2000]) to December 2017. For each database, the following search string was used: (depression OR major depressive disorder OR MDD OR major depression) AND (CBT OR cognitive behavi* therapy OR IPT OR interpersonal psychotherapy OR interpersonal therapy) AND (cohort OR longitudinal OR response OR panel OR prospective OR retrospective OR predictor).

Studies published in English were included if: (1) patients whose primary diagnosis was MDD (according to the DSM-III, DSM-III-R, DSM-IV, DSM-IV-TR, DSM-5 (APA, 2000),

ICD-9, or ICD-10 [WHO, 1992] criteria) (2) received individual or group CBT or IPT as the only psychological treatment of depression where (3) these therapies followed a recognised format and have not been altered or extended with other psychological components (e.g., integrated CBT); (4) were administered in a face-to-face setting, with (5) depression severity being quantitatively assessed both pre- and post-therapy using a validated measure with standardised cut-offs for mild, moderate, severe and very severe depression (e.g., HAM-D). Both randomised control trials and observational designs were included. Patients may have comorbidities, provided depression was the primary diagnosis. Antidepressant pharmacotherapy (ADMs) was permitted as the only adjunctive possible. Case series and online therapies were excluded.

Data Collection and Coding Procedures

References returned from each database were imported to the reference managing software Endnote X8 and duplicates removed. Two raters independently screened titles and abstracts of articles for potential inclusion. Disagreements were resolved through consensus discussions between the authors. Full texts of potentially eligible studies were then assessed against the inclusion criteria. Where sufficient data for inclusion was not reported (e.g., missing post-treatment depression severity score), a request for this information was e-mailed to the corresponding author.

From each study (*n*) who met the inclusion criteria, the following variables were coded for each included sample (*k*): (1) type of therapy administered, CBT or IPT; (2) number of participants at each assessment time point; (3) mean (*M*) and standard deviation (*SD*) of pre- and post-depression severity scores; (4) time delay (days) between pre- and post-treatment depression severity assessment: (5) patient demographics, including: mean age, gender and employment status; (6) clinical characteristics, including: initial depression severity, number of previous depressive episodes and comorbidities; and (7) therapy characteristics, including:

concomitant ADMs, inpatient or outpatient setting, group or individual format, and number of drop outs.

Statistical Methods/Meta-analyses

From each study, for each independent sample (either CBT or IPT), using the pre- and post-depression severity scores, standardised mean difference effect sizes (ES's) with confidence intervals (CI) were calculated for each sample using Cohen's *d* index of individual effects: $d_k = (M_{1k}-M_{2k})/SD_{pk}$, where *d* is the effect size, *k* the individual sample, M_{1k} pre-treatment mean, M_{2k} post-treatment mean and SD_{pk} is the pooled standard deviation A standardised ES is necessary as although all studies assessed depression severity, different scales were used. A positive ES indicated that depression severity was lower post-treatment. These ES's were interpreted according to Cohen's cut off recommendations of 1.30, 0.80, 0.50 and 0.20 for, respectively very large, large, medium and small ES (Cohen, 1988) .

CBT and IPT samples were then separately pooled using an inverse-variance weighted random-effects model on Comprehensive Meta Analysis software. The random-effects model is based on the assumption that different studies estimate different, yet related intervention effects, therefore assigning more weight to larger samples (DerSimonian & Laird, 1986). As studies pooled in this analysis varied in sample characteristics and depression scales administered heterogeneity will be substantial, thus justifying a random-effects model. Hedges' *g* correction for small sample bias was the effect size chosen for this software as it provides better estimates for smaller sample sizes.

Prior to conducting the moderator analyses, CBT and IPT were compared using subgroup analysis of variance. This analysis assumes between study variation allowing to determine if, overall, the therapies are comparable or significantly different in the treatment of depression

(Yusuff et al., 1991). Then, CBT and IPT samples were compared on the delay in days between pre-treatment and post-treatment to assess the need for controlling of this possible confound on the intervention effect. The I^2 index was also examined to further quantify true heterogeneity across the samples within each treatment not otherwise due to chance. This was interpreted using the recommended cut-offs: 25% small, 50% moderate and 75% large heterogeneity (Borenstein et al., 2009).

Moderator analyses were then conducted on the prespecified variables. Random effects method of moments meta-regressions were conducted when ≥ 10 studies reported on a continuous moderator variable (Borenstein et al., 2009). This analysis was used to demonstrate if the outcome variable (ES in depression severity change) is predicted by explanatory moderator variables across CBT and IPT therapies. Continuous moderators included; age, gender (% male), employed (%), number of previous episodes and number of dropouts. When a regression model was non-significant, but the moderator of interest was a significant predictor within the model, separate regressions were conducted for the CBT and IPT samples. If the model and the predictor were non-significant, no further analysis was conducted. Sub-group mixed effects meta-analyses were conducted when at least three studies per level were available for the categorical moderators, which included: initial depression severity (mild/moderate/severe/very severe); therapy format (individual/group); therapy setting (inpatient/ outpatient); comorbidities (yes /no); antidepressant medication (yes/ no). The categorical moderators were first compared at each level between CBT and IPT samples to determine if CBT or IPT were more effective in treating depression at that level. However, if the therapies were comparable and high levels of heterogeneity remain, then a within-therapy sub-group meta-analysis was also conducted to determine if, for that therapy, there was a significant moderator's effect on the main outcome.

To address publication bias, Orwin's Fail Safe *N* procedure was calculated to determine the

number of unpublished studies with null findings that would reduce our significant moderator findings to a 'trivial' ES's (Orwin, 1983). A trivial ES is defined as ≥ 0.18 (Cohen, 1988) . The likelihood of publication bias is minimal if the fail-safe *N* is > 5*k*+10 (Rosenthal, 1991). To control for potential Type-I errors associated with multiple comparisons, two steps were taken. Firstly, all moderators were chosen *a priori* with a clear rationale for selection (Bender et al., 2008). Secondly, the Holm's (1979) sequentially rejective multiple hypotheses test was applied. Adjustments for multiple testing are generally not recommended for meta-analyses (Bender et al., 2008) due to samples differing across outcome variables. However, we applied the correction for the family of tests at all levels of each moderator where such an overlap was possible (Bender et al., 2008; Higgins et al., 2008; Polanin, 2013; Striener, 2015). The Holm's correction combines the Bonferroni theorem with a step-up procedure. Initially, within each moderator, all *p*-values are ordered from smallest to largest. Beginning with the smallest, *p*-values were entered into the following formula : $a^* = a/(n-k)$. Whereby a^* represents the new alpha level, *a* is the original alpha level of 0.05, *n* is the number of tests and *k* is the position of the *p*-value in the ordered list. The p values were then tested against their new alpha level (a^*). As this is a step-up procedure once one null hypothesis is accepted ($p>a^*$) hypothesis testing stops (Polanin, 2013).

Results

After deleting duplicates, the search strategy identified 9,878 citations from which 758 were assessed for eligibility and 137 were meta-analysed. Figure 1 represents the review process.

INSERT FIGURE 1 ABOUT HERE

From the 137 studies (*n*), 168 samples (*k*) were extracted. A total of 11,374 patients (*N*), treated with either CBT *N*= 9,375 or IPT *N*= 1,999 were included in the analysis. Sample *N*s

ranged 7 to 639, with mean age of 38.34 for CBT (range 12.7-74.4) and 34.96 for IPT (10.6-51.2). There was no significant difference in mean ages between CBT and IPT samples, $t(158)=1.59$, $p=.11$, but there was a significant association between gender and therapy, $\chi^2(1)=76.53$, $p<.001$, with males being more likely to be treated with CBT: mean % males 34.6% than IPT: 25.5%. Also, no significant difference was observed on the pre-post treatment outcome assessment delay (days) between CBT ($M=102.17$, $SD=46.78$) and IPT ($M=103.38$, $SD=34.68$) samples, $t(135)=-0.146$, $p=.35$. Thus, even if spontaneous symptom recovery occurred within the samples, it has played equally in both CBT and IPT. Table 1 details percentage of meta-analysed studies reporting of each prespecified moderator. See full details on extracted data in Supplementary Tables (1 to 6).

INSERT TABLE 1 ABOUT HERE

Overall Treatment Efficacy

No significant differences were observed between the therapies with very large ES's characterising both CBT and IPT for the treatment of depression (Table 2).

INSERT TABLE 2 ABOUT HERE

Demographic Moderators

Age: Data from 160 samples were pooled to examine this moderator effect o: $k_{CBT}=114$ and $k_{IPT}=46$. No significant Q index was observed, $Q_{model}(2)=4.28$, $p=.12$. Age however, was a significant predictor in the model when controlling for therapy ($p=.044$). Thus, to further investigate its moderating effect, age was entered into separate regression models for CBT and IPT. In CBT samples, age was a significant predictor of efficacy, $Q_{model}(1)=5.21$, $p=.021$, with

CBT's efficacy decreasing as the mean age increased. Orwin's fail-safe *N* was robust, indicating that 950 unpublished studies with null effects would be needed to invalidate this result (benchmark$_{k=114}$ = 580). In IPT samples, age was not a significant predictor of efficacy (p=.53).

Gender: Data from 163 samples were pooled to examine this moderator effect : k_{CBT}= 117 and k_{IPT}=46. The model was non-significant $Q_{model}(2)$=0.91, p=.63. Gender was also not a significant predictor of treatment efficacy when controlling for therapy type (p=.40).

Employment status Data from 44 samples – k_{CBT}=30, k_{IPT}=14 – examined this moderator effect. *Q* index was not significant, $Q_{model}(2)$=0.54, p=.76 and employment remained a non-significant predictor of treatment efficacy when controlling for therapy (p=.50).

Clinical Moderators

Initial depression severity: Data from 152 samples – k_{CBT}=108, k_{IPT}=44 – were pooled to examine this moderator effect. No significant differences were observed between CBT and IPT in treating moderate (p=.93), severe (p=.19) or very severe depression (p=.81). All levels of initial depression significantly predicted treatment efficacy with very large ES's (p<.001). However, due to moderate to high levels of heterogeneity (I^2= 51.3-90.2%), within-therapy analyses were also conducted. For both CBT (k=3) and IPT (k=2), there were insufficient samples to include "very severe" level of initial depression in the subsequent moderator analyses. For CBT, sufficient samples were available to analyse this moderator effect for mild (*k=6),* moderate (k=65) or severe (k=34) initial depression. A significant difference was observed between levels of depression with efficacy increasing with initial depression severity (see Table 2). Orwin's fail-safe *N* was robust, indicating an additional 810 studies with null effects would be need to invalidate this result (benchmark$_{k=105}$ = 535). Post-hoc analyses were conducted to assess if CBT samples with severe initial depression had the greatest treatment effect due to potentially taking adjunctive ADM's. Post-hoc subgroup analysis of variance was

conducted on severe initial depression samples which reported adjunctive ADM's (*k*=13) or no ADM's (*k*=15). No significant difference was observed between the groups (*p*=.40), suggesting that CBT is significantly more effective for severe depression both alone and with ADM's. For IPT, sufficient samples were available to analyse this moderator effect for moderate (*k*=23) and severe (*k*=19) depression, with no significant difference being observed (*p*=.79).

Comorbidites: Data from 128 samples – k_{CBT}=94, k_{IPT}=34 – examined this moderator effect. Based on samples that reported percentages of comorbidities (k=49), on average 55.47% (range 10-100%) of participants presented with at least one comorbid Axis I or II disorder. No significant difference was observed between CBT or IPT efficacy whether comorbidities were present (*p*=.38) or not (*p*=.68). High heterogeneity was observed in both subgroups justifying within-therapy analysis. No significant difference in efficacy for depression with or without comorbidities were observed neither within the CBT samples (*p*=.12), nor within the IPT samples (*p*=.16).

Number of previous episodes: Insufficient number of IPT samples were available to analyse this moderator effect (k=5). Meta-regression was thus carried out on CBT samples, k=20, only and led to a non-significant *Q* index for number of previous episodes as a predictor of CBT efficacy $Q_{model}(1)=0.00$, *p*=.997.

Therapy Moderators

Format: Data from 141 samples – k_{CBT}=105, k_{IPT}=36 – examined this moderator effect. No significant differences were found between IPT and CBT delivered in a group (*p*=.52) or individual (*p*=.19) format. Both formats significantly predicted treatment efficacy (*p*<.001) with very large ES's. Heterogeneity was high in both sub-group justifying within-therapy analysis. For CBT, a significant difference was observed between the two formats: *Q*(1)=10.75, *p*=.001, with individual therapy (*g*=1.72, 95%CI[1.55, 1.90], *k*=81) showing better efficacy than group therapy (*g*= 1.31, 95%CI[1.12, 1.49]; *k*=24). Orwin's fail-safe *N* indicated an

additional 786 samples with null effects would be need to invalidate this result (benchmark$_{k=105}$ = 535). For IPT, there was no significant difference between individual (k=30) and group (k=6) formats: Q(1)= 0.003, p=.96, with both displaying positive very large ES's (p<.001).

Setting: Insufficient data were available to analyse inpatient setting for IPT (k=1). Data from 156 samples – k_{CBT}=112, k_{IPT}=44 – examined the relative efficacy of the two treatments in an outpatient setting. No significant difference was observed: p=.32. For CBT, there was no significant difference between inpatient (k=6) and outpatient (k=112) settings: p=.26, with both settings being associated with positive very large ES's (p<.001).

Concomitant ADM's: Data from 141 samples – k_{CBT}=101, k_{IPT}=40 – examined this effect. No significant between-treatment differences were found between IPT and CBT for concomitant ADM's. ADM's prescribed were SSRI's (e.g. citalopram), SNRI's (e.g. venlafaxine) and Tricyclic (e.g. amitriptyline) (See Supplementary Tables 5 and 6). For samples without concomitant ADM's, CBT (g=1.82, 95%CI[1.63, 2.05]) was associated with significantly better efficacy than IPT (g=1.54, 95%CI[1.34, 1.73]), p=.037. Orwin's fail-safe N indicated an additional 700 samples with null effects would be need to invalidate this result (benchmark$_{k=84}$ = 430). In CBT samples, absence to concomitant ADM's led to larger clinical improvements (g=1.82, 95%CI[1.63, 2.01]) than concomitant ADM's (g=1.42, 95%CI[1.24, 1.61]), Q(1)=8.74, p=.003. Orwin's fail-safe N indicated an additional 808 samples with null effects would be needed to invalidate this result (benchmark$_{k=101}$ = 515). For IPT, no significant difference was observed between receiving concomitant ADMs or not (p=.56), with both prescriptions leading to positive very large ES's (p<.001).

Number of dropouts: Data from 107 samples were pooled to examine the moderator effect of number of dropouts: k_{CBT}=77 and k_{IPT}=30. No significant Q index was observed Q_{model}(2)=2.50, p=.29. The number of dropouts was also not a significant predictor of efficacy when controlling for therapy type (p=.16).

***Post hoc* sensitivity analyses of study's design effects**

Post-hoc sensitivity analyses were conducted to determine if study design, that is RCTs versus non-RCT studies, has affected the results. These were conducted on all moderators and across all levels of moderators. All analyses, except two relating to the therapy moderator ADMs, showed equivalent results between RCTs and non-RCTs. Specifically, the overall treatment effects were similar both within CBT (g_{RCT}=1.72, $g_{non\text{-}RCT}$=1.52, p=0.13) and IPT samples (g_{RCT}=1.60, $g_{non\text{-}RCT}$=1.52, p=0.67). There were no significant effects of design on any demographic (all p-values>0.095) or clinical (all p-values>0.079) moderators. Interestingly, with the exception of studies where only patients with no comorbidity were studied, for both therapies, RCTs tended to show larger, although not significantly so, effect sizes of effectiveness compared to non-RCTs. With regards to concomitant ADMs, *post-hoc* analyses showed that, overall, RCTs led to significantly larger effect sizes (g=1.83, CI [1.52, 2.14]) than non-RCT's (g=1.17, CI [0.99, 1.34]), p<.001). This result was explained by a therapy effect: there was not a significant difference between RCTs and non-RCTs for IPT (p=0.19), but there was for CBT (g_{RCT}=1.80 CI[1.44, 2.15]; $g_{non\text{-}RCT}$=1.10, CI[1.70,2.25], p=0.001). As an effect for design was found for this moderator, for RCT with concomitant ADMs only, CBT and IPT were compared, but no significant difference was observed (p=0.72). See Supplementary Table 7 for full results of sensitivity analyses.

Results following Holm's (1979) sequentially rejective multiple hypotheses test

Holm's (1979) test was applied where overlapping samples were observed. Among these, three out of five significant p-values remained significant: Namely, the findings that CBT is significantly moderated by age (p=.021), therapy format (p=0.001) and antidepressant medications (p=0.003) remained significant under their respective adjusted alpha levels within their moderator's family of tests. However, the result that CBT was significantly more effective than IPT without concomitant antidepressants (p=0.037) did not remain significant after applying a more conservative alpha of 0.025. Similarly, the significant moderating effect of initial depression on CBT (p=0.022) did not remain significant under the adjusted alpha of

0.0125. For a full overview of the Holm's procedure and all adjusted alpha levels see Supplementary Table 8.

Discussion

The current systematic review and meta-analyses aimed at determining moderators of efficacy of face-to-face CBT and IPT for major depressive disorder. Our results further supported existing evidence of the equivalent overall treatment effects of CBT and IPT for depression (Miranda et al., 2005; Cuijpers et al., 2006; Weisz et al., 2006). Between-therapy moderator analyses also showed comparable efficacy of CBT and IPT across age, gender, employment status, initial depression severity, presence of comorbidities, number of previous episodes, therapy formats, therapy settings and number of dropouts. However, a significant difference between CBT and IPT was observed when examining the prespecified moderator of concomitant ADM's use. Specifically, CBT was superior to IPT in treating depression when therapies were administered alone, i.e., without adjunctive ADM's.

Within-therapy analyses showed the effect of CBT to be moderated by age, initial depression severity, therapy format and adjunctive ADM's. Namely, CBT's efficacy declined as patients' age increased and was more effective in treating severe initial depression than moderate or mild depression. CBT was also more effective when delivered in an individual rather than in group format and when administered alone rather than with concomitant ADM's. Within-therapy analyses of IPT did not identify any significant effect on efficacy of the preselected moderators.

CBT's efficacy moderators

Few individual studies have compared CBT efficacy across age groups. However, the present meta-analysis, cumulating strong power through analyses of 9 375 participants

from120 samples, shows decline of efficacy with sample's mean age increase. This result is consistent with trial findings showing that increasing age predicted poorer response to cognitive therapy (Fournier et al., 2009). Other researchers have also expressed concern over the use of CBT with older patients, as they consider the cognitive slowing associated with aging to negatively impact the treatment delivery (Hyer et al., 2004). Indeed, CBT staples of assigning homework or challenging distorted cognitions, both outlined as less effective and less preferred for older patients (Hyer et al., 2004). Considering age as a prescriptive factor has important clinical applications for both treatment selection and in the development of personalised treatment guidelines with regards to age.

The moderation of CBT efficacy by the therapy format supports and expands on the results of an earlier meta-analysis of 15 studies suggesting that individual CBT might be more effective than group CBT based on post-treatment depression scores alone (Cuijpers et al., 2008). At the time, the authors recommended further research given their sample size. Our meta-analysis of 8,004 participants drawn from 105 samples demonstrates the significant superiority of individual CBT relative to group CBT with very strong power and based on comparisons of differences between pre- and post-treatment changes. Possible explanations for the benefit of individual therapies centre around the nature of CBT for depression. Due to the severity of depression and associated symptoms, patients may find it easier to engage with CBT in an individual setting (Craigie et al., 2009). This finding has important clinical implications as currently, group formats are widely disseminated, partly due to their apparent cost-effectiveness. Although group formats are still effective in reducing depressive symptoms, individual CBT is a significantly more efficacious therapy, supporting further the importance of treatment personalisation.

Drawing strong power from the analysis of 7163 participants from 105 samples, results demonstrated that CBT was significantly more effective for those with severe depression in comparison to moderate or mild depression, even when controlling for adjunctive ADM's.

Although this finding did not remain significant under the Holm's (1979) multiple testing correction, it is important to note that corrections of multiplicity are not routinely used in meta-analysis as their test assumptions are rarely met by meta-analytic data. Therefore, multiple testing correction results within this study should be interpreted with caution (Bender et al. 2008; Higgins et al. 2008; Streiner 2015). Nevertheless, finding that CBT is more effective for severe depression is consistent with the results of a naturalistic study of 193 patients with depression where this superiority was partly attributed to a regression of the mean, with depression scores that are significantly higher than the mean (high severity) pre-treatment being likely to become closer to the mean at re-assessment assessment (Schindler et al., 2013). While this may be a possibility, the effect sizes observed in the present meta-analysis were significantly different at all severity levels, making it unlikely that a regression to the mean occurred. From a clinical perspective, severity of depression appears as an important prescriptive factor for treatment personalisation. Patients with severe depression are often prescribed ADM's alone in line with practitioners' guidelines (NICE, 2009). The present results suggest that CBT should be consistently considered as a first-line treatment for severe depression.

Furthermore, in relation to ADM's use, our meta-analysis of 8,421 participants drawn from 101 samples showed that CBT alone, with no ADM's, was significantly more effective than CBT with concomitant ADMs. This result is at odds with some clinical guidelines that recommend the use of CBT plus ADM's (NICE, 2009). However, this result does align with both a narrative review and a meta-analysis suggesting that CBT, unlike other psychotherapies, is much less effective in combination with ADMs than alone (Cuijpers et al., 2009; Craighead et al., 2014). One possible explanation for this finding may be acceptance of ADM's, depressed individuals are three-times more likely to choose psychological therapies over ADM's (McHugh et al., 2013). While a possibility, it is unlikely this explains the current finding as most included samples allowed patients to continue their previous ADM treatment instead of

newly prescribing ADM's. However, future studies might assess this hypothesis. Another, more plausible explanation for this finding is related to Johnstone's (2003) argument that ADM's can potentially limit CBT engagement. While relieving mood symptoms, ADMs trigger an emotional blunting which conflicts with the very nature of CBT (Fava et al., 2018). Their withdrawal symptoms and common side-effects of anxiety, insomnia, and agitation can further hinder therapy process and engagement (Churchill et al., 2000; Johnstone, 2003; Fava, 2003; Butler et al., 2006; Bockting et al., 2018, Fava et al., 2018). The meta-analysis results suggest that prescribing ADM's alongside CBT in future clinical practice should be considered on a careful, case-by-case basis as the therapy alone may prove more effective for the treatment of mild, moderate and severe depression in comparison to combined with adjunctive ADM's treatment.

Moderators of IPT efficacy

IPT showed equivalent treatment effects across all prespecified moderators. This contrasts with previous studies outlining IPT as less effective in the presence of comorbidities (Cyranowski et al., 2005); or for high-severity depression (Frank et al., 2011). A possible explanation for this finding can be related to sample sizes – 18 (Cyranowski et al., 2005) and 117 (Frank et al., 2011) patients received IPT in these studies, thus limiting their respective results' generalizability. Our large IPT sample (N=1,999) allows the conclusion that IPT appears to be equally effective for the treatment of depression across these moderators.

Despite IPT's comparable efficacy to CBT, it remains a much less prescribed treatment for depression. Only one third of samples retrieved received IPT. This represents a limitation of the meta-analysis, as insufficient data was available to investigate if the number of previous episodes, or very severe depression, moderated the treatment outcome. Similarly, although data were sufficient to examine age as moderator of IPT outcomes, these samples had a maximum mean age of 51, unlike CBT samples with a maximum mean age of 74. Nevertheless, IPT was shown an effective treatment for depression across a range of demographic, clinical and

therapeutic moderators. IPT should therefore be consistently considered in clinical guidelines, applied more frequently in clinical practice and further investigated for other moderators of efficacy.

Limitations of Within-Therapy Analysis

Firstly, high levels of heterogeneity were observed throughout the analysis. One possible reason is that the label CBT has been applied to a variety of interventions not always reflect the pure therapy strain. Although inclusion criteria aimed to control for this by specifying CBT must not to be altered or extended, in practice, as the therapy is often modified beyond focus, this might not always be reported in the original articles (Hyer et al., 2004). Secondly, even though significant moderator effects were strong, they could not account for most of the variance in treatment outcomes. These moderators may be interacting to explain variability, but complex interactive models were not examined, as larger samples would be needed for reliable results. Thirdly, recovery or remission rates were not analysed, thus limiting the current results to post-treatment outcomes relative to the pre-treatment depression severity. However, by optimising in this way the overall sample analyses, the meta-analytical findings are representative of existing research variations and have stronger power than selective (non-representative) sampling.

Between-Therapies Comparisons Everyone Has Won, all must Get Prizes?

In line with previous meta-analytic studies, our research demonstrated that overall, CBT and IPT are equally effective in treating depression, therefore supporting again the Dodo bird verdict (Miranda et al. 2005; Cuijpers et al. 2006; Weisz et al. 2006). This verdict also spread wings across all but one of the preselected ten moderators. Interestingly, the meta-analysis showed that CBT was significantly more effective than IPT for the treatment of depression when there is no concomitant ADM's prior to a multiple testing correction. Therefore, considering (a) the significant difference between CBT and IPT when ADM's were excluded

and (b) the fact that previous intervention studies displayed mixed findings, can we really conclude from this analysis that in the end, everyone has won, all must have prizes?

One possible explanation for this finding is that when controlling for ADM's, CBT is in fact more effective than IPT. While this contradicts the Dodo bird verdict and previous meta-analytic research, it is not spurious. For example, returning to the aforementioned trials by Luty et al. (2007) and Rossello et al. (2008) both excluded patients on ADM's and came to the same conclusion that CBT was more effective than IPT. While this explanation is possible, high levels of heterogeneity remained and this result did not remain significant under a conservative alpha correction; therefore, again, interaction effects may be occurring. A second explanation can be related back to the within-therapy findings on CBT and ADM's. While the current meta-analysis supports previous research conclusions that combined treatment may be overvalued and not necessarily required for CBT, the same has not been demonstrated for IPT (Hollon et al., 2005; Mintz et al., 2006). As a result, CBT without ADM's may be superior compared to both CBT plus ADM's and IPT without ADM's, which has important clinical implications. The application of CBT alone at most severity levels will not only contribute to favourable costs, it will also be a better long-term treatment plan avoiding the aforementioned deleterious effects of long-term ADM use (Fava et al., 2003; Butler et al., 2006; Otto et al., 2006).

Excluding the effect of concomitant ADM therapy, CBT and IPT demonstrated equal efficacy across all other moderators. Unlike the most recent meta-analysis conducted by Zhou et al. (2017) and the majority of single moderator studies, our meta-analysis was not limited to RCT's, as recommended by Westen and Novotny (2004). While this increased the external validity and generalizability of the current results, as any meta-analysis, ours was also constrained by the studies included (Harrison, 2011). Closer inspection of the moderator 'comorbidities' demonstrated that almost half of the studies completely excluded patients who presented with comorbidities. Even though this was not the case for all, according to the

average percentage of participants presenting with comorbidities in this study (55.47%), this exclusion of comorbidities resulted in potentially half of the depressed population missing from many analyses, thus perhaps interfering with the identification of some other significant moderators. Nevertheless, considering that our significant findings were robust against publication bias assessment and the sufficient power to analyse all prespecified moderators, our findings strongly support the comparable short-term efficacy of face-to-face CBT and IPT, while highlighting the importance of moderating variables within each therapy.

Our meta-analysis also has important clinical and research implications. Firstly, researchers may need to reconsider existing clinical guidelines asking for the examination of efficacy moderators by addressing the factors that might prevent us from identifying useful moderators. Future research should consider depression *together* with all its complexities and focus on more ecological definition of the disorder. If studies continue to exclude patients with comorbidities, whom represent the MDD reality, actual therapy efficacy will never be comprehensively examined and thus conclusions on relapse will remain lacking. Secondly, cumulating the present findings, this study also suggests that combined therapy (psychotherapy with concomitant ADM's) should be re-considered. Such treatment is currently more expensive and shows little evidence of superiority for MDD. Thus, considering the superiority of CBT alone and the side-effects, tapering problems and withdrawal symptoms associated with ADM's, combined treatment should be prescribed carefully, only in complex cases and on a case-by-case basis (Bockting et al., 2018; Fava et al., 2018). Therefore, supporting the conclusions of a recent meta-analysis conducted by Fava and colleagues (2018), the use of ADM's, in this case combined with therapy, should only be targeted at the most persistent cases of MDD and for the shortest possible time.

References

APA (2000) *Diagnostic and statistical manual of mental disorders* 4th ed., Washington DC: American Psychiatric Association.

Barth, J., Munder, T., Gerger, H., Nuesch, E., Trelle, S., Znoj, H., Juni, P. and Cuijpers, P. (2013) 'Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis', *PLoS Med*, 10(5), e1001454, available: http://dx.doi.org/10.1371/journal.pmed.1001454.

Bayliss, P. and Holttum, S. (2015) 'Experiences of antidepressant medication and cognitive-behavioural therapy for depression: a grounded theory study', *Psychology Psychotherapy*, 88(3), 317-34, available: http://dx.doi.org/10.1111/papt.12040.

Bockting, C.L.H., Klein, N.S., Elgersma, H.J., van Rijsbergen, G.D., Slofstra, C., Ormel, J., Buskens, E., Dekker, J., de Jong, P.J., Nolen, W.A., Schene, A.H., Hollon, S.D. and Burger, H. (2018) 'Effectiveness of preventive cognitive therapy while tapering antidepressants versus maintenance antidepressant treatment versus their combination in prevention of depressive relapse or recurrence (DRD study): a three-group, multicentre, randomised controlled trial', *The Lancet Psychiatry*, 5(5), 401-410, available: http://dx.doi.org/https://doi.org/10.1016/S2215-0366(18)30100-7.

Bender, R., Bunce, C., Clarke, M., Gates, S., Lange, S., Pace, N.L. and Thorlund, K.,(2008). Attention should be given to multiplicity issues in systematic reviews. Journal of clinical epidemiology, 61(9), pp.857-865.

Benjamini, Y. and Hochberg, Y., (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), pp.289-300

Borenstein, M., Hedges, L., Higgins, J. and Rothstein, H. (2009) *Introduction to meta-analysis*, West Sussex, England: Wiley.

Budd, R. and Hughes, I. (2009) 'The Dodo Bird Verdict--controversial, inevitable and important: a commentary on 30 years of meta-analyses', *Clin Psychology and Psychotherapy*, 16(6), 510-22, available: http://dx.doi.org/10.1002/cpp.648.

Burcusa, S.L. and Iacono, W.G. (2007) 'Risk for recurrence in depression', *Clinical Psychology Review*, 27(8), 959-985, available: http://dx.doi.org/10.1016/j.cpr.2007.02.005.

Butler, A.C., Chapman, J.E., Forman, E.M. and Beck, A.T. (2006) 'The empirical status of cognitive-behavioral therapy: a review of meta-analyses', *Clinical Psychology Review*, 26(1), 17-31, available: http://dx.doi.org/10.1016/j.cpr.2005.07.003.

Carter, J.D., Luty, S.E., McKenzie, J.M., Mulder, R.T., Frampton, C.M. and Joyce, P.R. (2011) 'Patient predictors of response to cognitive behaviour therapy and interpersonal psychotherapy in a randomised clinical trial for depression', *Journal of Affective Disord*ers, 128(3), 252-61, available: http://dx.doi.org/10.1016/j.jad.2010.07.002.

Churchill, R., Khaira, M., Gretton, V., Chilvers, C., Dewey, M., Duggan, C., Lee, A., Nottingham, C. and Antidepressants in Primary Care Study, G. (2000) 'Treating depression in general practice: factors affecting patients' treatment preferences', *The British journal of general practice : the journal of the Royal College of General Practitioners*, 50(460), 905-906.

Cohen, J.D. (1988) *Statistical power analysis for the behavioural sciences.*, Hillsdale, NJ: Erlbaum.

Craighead, W.E. and Dunlop, B.W. (2014) 'Combination psychotherapy and antidepressant medication treatment for depression: for whom, when, and how', *Annual Review Psychology*, 65, 267-300, available: http://dx.doi.org/10.1146/annurev.psych.121208.131653.

Craigie, M.A. and Nathan, P. (2009) 'A nonrandomized effectiveness comparison of broad-spectrum group CBT to individual CBT for depressed outpatients in a community mental health setting', *Behavior Therapy*, 40(3), 302-14, available: http://dx.doi.org/10.1016/j.beth.2008.08.002.

Cuijpers, P., Andersson, G., Donker, T. and van Straten, A. (2011) 'Psychological treatment of depression: results of a series of meta-analyses', *Nordic Journal of Psychiatry*, 65(6), 354-64, available: http://dx.doi.org/10.3109/08039488.2011.596570.

Cuijpers, P., Ebert, D.D., Acarturk, C., Andersson, G. and Cristea, I.A. (2016) 'Personalized Psychotherapy for Adult Depression: A Meta-Analytic Review', *Behaviour Therapy*, 47(6), 966-980, available: http://dx.doi.org/10.1016/j.beth.2016.04.007.

Cuijpers, P., van Straten, A. and Smit, F. (2006) 'Psychological treatment of late-life depression: a meta-analysis of randomized controlled trials', *International Journal of Geriatric Psychiatry*, 21(12), 1139-1149, available: http://dx.doi.org/doi:10.1002/gps.1620.

Cuijpers, P., van Straten, A. and Warmerdam, L. (2008) *Are individual and group treatments equally effective in the treatment of depression in adults? A meta-analysis* [press release], available: [accessed

Cuijpers, P., van Straten, A., Warmerdam, L. and Andersson, G. (2009) 'Psychotherapy versus the combination of psychotherapy and pharmacotherapy in the treatment of depression: a meta-analysis', *Depression and Anxiety*, 26(3), 279-88, available: http://dx.doi.org/10.1002/da.20519.

Cyranowski, J.M., Frank, E., Shear, M.K., Swartz, H., Fagiolini, A., Scott, J. and Kupfer, D.J. (2005) 'Interpersonal Psychotherapy for Depression With Panic Spectrum Symptoms: A Pilot Study', *Depression and Anxiety*, 21(3), 140-142, available: http://dx.doi.org/10.1002/da.20069.

DerSimonian, R. and Laird, N., (1986). Meta-analysis in clinical trials. *Controlled clinical trials*, *7*(3), pp.177-188.

Driessen, E., Smits, N., Dekker, J.J., Peen, J., Don, F.J., Kool, S., Westra, D., Hendriksen, M., Cuijpers, P. and Van, H.L. (2016) 'Differential efficacy of cognitive behavioral therapy and psychodynamic therapy for major depression: a study of prescriptive factors', *Psychological Medicine*, 46(4), 731-44, available: http://dx.doi.org/10.1017/s0033291715001853.

Elkin, I., Gibbons, R.D., Shea, M.T., Sotsky, S.M., Watkins, J.T., Pilkonis, P.A. and Hedeker, D. (1995) 'Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program', *Journal of Consulting and Clinical Psychology*, 63(5), 841-7.

Fava, G.A. (2003) 'Can long-term treatment with antidepressant drugs worsen the course of depression?', *Journal of Clinical Psychiatry*, 64(2), 123-33.

Fava, G.A., Benasi, G., Lucente, M., Offidani, E., Cosci, F. and Guidi, J. (2018) 'Withdrawal Symptoms after Serotonin-Noradrenaline Reuptake Inhibitor Discontinuation: Systematic Review', *Psychotherapy and Psychosomatics*, 87(4), 195-203, available: http://dx.doi.org/10.1159/000491524.

Fournier, J.C., DeRubeis, R.J., Shelton, R.C., Hollon, S.D., Amsterdam, J.D. and Gallop, R. (2009) 'Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression', *Journal of Consulting and Clinical Psychology*, 77(4), 775-87, available: http://dx.doi.org/10.1037/a0015401.

Frank, E., Cassano, G.B., Rucci, P., Thompson, W.K., Kraemer, H.C., Fagiolini, A., Maggi, L., Kupfer, D.J., Shear, M.K., Houck, P.R., Calugi, S., Grochocinski, V.J., Scocco, P., Buttenfield, J. and Forgione, R.N. (2011) 'Predictors and moderators of time to remission of major depression with interpersonal psychotherapy and SSRI pharmacotherapy', *Psychological Medicine*, 41(1), 151-162, available: http://dx.doi.org/10.1017/S0033291710000553.

Gartlehner, G. and Moore, C.G. (2008) 'Direct versus indirect comparisons: A summary of the evidence', *International Journal of Technology Assessment in Health Care*, 24(2), 170-177, available: http://dx.doi.org/10.1017/S0266462308080240.

Harrison, F. (2011) 'Getting started with meta-analysis', *Methods in Ecology and Evolution*, 2(1), 1-10, available: http://dx.doi.org/10.1111/j.2041-210X.2010.00056.x.

Higgins, J.P., Deeks, J.J. and Altman, D.G., (2008.) Special topics in statistics. *Cochrane handbook for systematic reviews of interventions: Cochrane book series*, pp.481-529.

Hofmann, S.G., Asnaani, A., Vonk, I.J., Sawyer, A.T. and Fang, A. (2012) 'The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses', *Cognitive Therapy Research*, 36(5), 427-440, available: http://dx.doi.org/10.1007/s10608-012-9476-1.

Hollon, S.D., DeRubeis, R.J., Shelton, R.C., Amsterdam, J.D., Salomon, R.M., O'Reardon, J.P., Lovett, M.L., Young, P.R., Haman, K.L., Freeman, B.B. and Gallop, R. (2005) 'Prevention of relapse following cognitive therapy vs medications in moderate to severe depression', *Archives of General Psychiatry*, 62(4), 417-22, available: http://dx.doi.org/10.1001/archpsyc.62.4.417.

Hollon, S.D., Jarrett, R.B., Nierenberg, A.A., Thase, M.E., Trivedi, M. and Rush, A.J. (2005) 'Psychotherapy and medication in the treatment of adult and geriatric depression: which monotherapy or combined treatment?', *Journal of Clinical Psychiatry*, 66(4), 455-68.

Holm, S., (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, pp.65-70.

Honyashiki, M., Furukawa, T.A., Noma, H., Tanaka, S., Chen, P., Ichikawa, K., Ono, M., Churchill, R., Hunot, V. and Caldwell, D.M. (2014) 'Specificity of CBT for Depression: A Contribution from Multiple Treatments Meta-analyses', *Cognit Therapy Research*, 38(3), 249-260, available: http://dx.doi.org/10.1007/s10608-014-9599-7.

Hyer, L., Kramer, D. and Sohnle, S. (2004) 'CBT With Older People: Alterations and the Value of the Therapeutic Alliance', *Psychotherapy: Theory, Research, Practice, Training*, 41(3), 276-291, available: http://dx.doi.org/10.1037/0033-3204.41.3.276.

Johnstone, L. (2003) 'Back to Basics', *The Psychologist*, (16), 186-187.

Kessler, R.C., Nelson, C.B., McGonagle, K.A., Liu, J., Swartz, M. and Blazer, D.G. (1996) 'Comorbidity of DSM-III-R major depressive disorder in the general population: results from

the US National Comorbidity Survey', *British Journal of Psychiatry Supplement*, 1996/06/01(30), 17-30.

Law, R. (2011) 'Interpersonal psychotherapy for depression', *Advances in Psychiatric Treatment*, 17(1), 23-31, available: http://dx.doi.org/10.1192/apt.bp.109.007641.

Lemmens, L.H., Arntz, A., Peeters, F.P., Hollon, S.D., Roefs, A. and Huibers, M.J. (2011) 'Effectiveness, relapse prevention and mechanisms of change of cognitive therapy vs. interpersonal therapy for depression: Study protocol for a randomised controlled trial', *Trials*, 12, 150, available: http://dx.doi.org/10.1186/1745-6215-12-150.

Luty, S.E., Carter, J.D., McKenzie, J.M., Rae, A.M., Frampton, C.M., Mulder, R.T. and Joyce, P.R. (2007) 'Randomised controlled trial of interpersonal psychotherapy and cognitive-behavioural therapy for depression', *British Journal of Psychiatry*, 190, 496-502, available: http://dx.doi.org/10.1192/bjp.bp.106.024729.

McBride, C., Atkinson, L., Quilty, L.C. and Bagby, R.M. (2006) 'Attachment as moderator of treatment outcome in major depression: a randomized control trial of interpersonal psychotherapy versus cognitive behavior therapy', *Journal of Consulting and Clinical Psychology*, 74(6), 1041-54, available: http://dx.doi.org/10.1037/0022-006x.74.6.1041.

McHugh, R.K., Whitton, S.W., Peckham, A.D., Welge, J.A. and Otto, M.W. (2013) 'Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: a meta-analytic review', *Journal of Clinical Psychiatry*, 74(6), 595-602, available: http://dx.doi.org/10.4088/JCP.12r07757.

Mintz, D. (2006) 'Combining drug therapy and psychotherapy for depression', *Psychiatric Times*, 23(11), 18.

Miranda, J., Bernal, G., Lau, A., Kohn, L., Hwang, W.-C. and LaFromboise, T. (2005) 'State of the science on psychosocial interventions for ethnic minorities', *Annual review of clinical psychology*, 1, 113-142, available: http://dx.doi.org/10.1146/annurev.clinpsy.1.102803.143822.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G. and Group, a.t.P. (2009) 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA StatementThe PRISMA Statement', *Annals of Internal Medicine*, 151(4), 264-269, available: http://dx.doi.org/10.7326/0003-4819-151-4-200908180-00135.

NICE (2009) *The treatment and management of depression in adults (CG90)*: National Institute for Health and Clinical Excellence.

Orwin, R.G. (1983) 'A fail-safe N for effect size in meta-analysis', 8, 157-159, available: http://dx.doi.org/10.2307/1164923.

Otto, M.W., Smits, J.A.J. and Reese, H.E. (2005) 'Combined Psychotherapy and Pharmacotherapy for Mood and Anxiety Disorders in Adults: Review and Analysis', *Clinical Psychology: Science and Practice*, 12(1), 72-86, available: http://dx.doi.org/10.1093/clipsy.bpi009.

Polanin, J.R., (2013). Addressing the issue of meta-analysis multiplicity in education and psychology.

Rosenthal, R. (1991) 'Meta-analysis: A review', *Psychosomatic Medicine*, 53(3), 247-271, available: http://dx.doi.org/10.1097/00006842-199105000-00001

Rossello, J., Bernal, G. and Rivera-Medina, C. (2008) 'Individual and group CBT and IPT for Puerto Rican adolescents with depressive symptoms', *Cultural Diversity and Ethnic Minority Psychology*, 14(3), 234-45, available: http://dx.doi.org/10.1037/1099-9809.14.3.234.

Rucci, P., Frank, E., Calugi, S., Miniati, M., Benvenuti, A., Wallace, M., Fagiolini, A., Maggi, L., Kupfer, D.J. and Cassano, G.B. (2011) 'Incidence and predictors of relapse during continuation treatment of major depression with SSRI, interpersonal psychotherapy, or their combination', *Depression and Anxiety*, 28(11), 955-62, available: http://dx.doi.org/10.1002/da.20894.

Schindler, A., Hiller, W. and Witthoft, M. (2013) 'What predicts outcome, response, and drop-out in CBT of depressive adults? a naturalistic study', *Behavior Cognitive Psychotherapy*, 41(3), 365-70, available: http://dx.doi.org/10.1017/s1352465812001063.

Shapiro, D.A., Barkham, M., Rees, A., Hardy, G.E., Reynolds, S. and Startup, M. (1994) 'Effects of treatment duration and severity of depression on the effectiveness of cognitive-behavioral and psychodynamic-interpersonal psychotherapy', *Journal of Consulting and Clinical Psychology*, 62(3), 522-34.

Shinohara, K., Honyashiki, M., Imai, H., Hunot, V., Caldwell, D.M., Davies, P., Moore, T.H., Furukawa, T.A. and Churchill, R. (2013) 'Behavioural therapies versus other psychological therapies for depression', *Cochrane Database Systematic Reviews*, 2013/10/17(10), Cd008696, available: http://dx.doi.org/10.1002/14651858.CD008696.pub2.

Streiner, D.L., (2015). Best (but oft-forgotten) practices: the multiple problems of multiplicity—whether and how to correct for many statistical tests. *The American journal of clinical nutrition*, *102*(4), pp.721-728.

Thase, M.E., Greenhouse, J.B., Frank, E., Reynolds, C.F., 3rd, Pilkonis, P.A., Hurley, K., Grochocinski, V. and Kupfer, D.J. (1997) 'Treatment of major depression with psychotherapy or psychotherapy-pharmacotherapy combinations', *Archive of General Psychiatry*, 54(11), 1009-15.

Thase, M.E., Reynolds, C.F., Frank, E., Simons, A.D., McGeary, J., Fasiczka, A.L., Garamoni, G.G., Jennings, J.R. and Kupfer, D.J. (1994) *Do depressed men and women respond similarly to cognitive behavior therapy?* [press release], available: [accessed

Weisz, J.R., McCarty, C.A. and Valeri, S.M. (2006) 'Effects of psychotherapy for depression in children and adolescents: a meta-analysis', *Psychology Bulletin*, 132(1), 132-149, available: http://dx.doi.org/10.1037/0033-2909.132.1.132.

Westen, D., Novotny, C.M. and Thompson-Brenner, H. (2004) 'The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials', *Psychology Bulletin*, 130(4), 631-63, available: http://dx.doi.org/10.1037/0033-2909.130.4.631.

WHO (1992) 'The ICD-10 Classification of Mental and Behavioral Disorders'.

WHO (2017) *Depression Fact Sheet*, available: http://www.who.int/news-room/fact-sheets/detail/depression [accessed

Yusuf, S., Wittes, J., Probstfield, J. and Tyroler, H.A., (1991). Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *Jama*, *266*(1), pp.93-98.

Zhou, S.-G., Hou, Y.-F., Liu, D. and Zhang, X.-Y. (2017) 'Effect of Cognitive Behavioral Therapy Versus Interpersonal Psychotherapy in Patients with Major Depressive Disorder: A Meta-analysis of Randomized Controlled Trials', *Chinese Medical Journal*, 130(23), 2844-2851, available: http://dx.doi.org/10.4103/0366-6999.219149.

Table 1. Percentage of moderators reported across all CBT and IPT samples.

	% Reported	
Moderator	**CBT**	**IPT**
Demographic		
Age	95.00	95.83
Gender	97.50	95.83
Employment	25.00	29.17
Clinical		
Initial depression severity	87.50	87.50
Comorbidities	78.33	64.58
Number of previous episodes	16.67	11.90
Therapy		
Format	87.50	75.00
Setting	98.33	93.75
Antidepressant medications	84.17	83.33
Dropouts	64.17	62.50

Table 2. Meta-analyses of overall treatment effect and moderator effects of initial depression severity, comorbidities, therapy format, therapy setting and concomitant antidepressant medications

				95%CI				
Therapy	***k***	***N***	***g***	***LL***	***UL***	***Zw***	***Qb***	***p***
Overall Treatment Efficacy								
CBT	120	9,375	1.63	1.50	1.76	25.22		
IPT	48	1,999	1.58	1.42	1.74	19.33	0.26	0.610
Initial Depression Severity								
CBT V IPT								
Moderate	88	5,486	1.57	1.43	1.72	21.69	0.007	0.932
Severe	53	2,648	1.62	1.47	1.77	21.16	1.75	0.185
Very Severe	5	353	1.16	0.94	1.38	10.26	0.06	0.813
CBT								
Mild	6	815	1.14	0.78	1.49	6.29		
Moderate	65	4,555	1.58	1.39	1.75	17.06		
Severe	34	1,793	1.72	1.44	1.70	15.27	7.65	0.022
IPT								
Moderate	23	931	1.57	1.34	1.79	13.41		
Severe	19	855	1.52	1.32	1.73	14.71	0.07	0.789
Comorbidities								
CBT V IPT								
Yes	71	6,241	1.59	1.44	1.73	22.04	0.77	0.379
No	57	2,520	1.85	1.66	2.05	18.68	0.18	0.676
CBT								
Yes	50	5,174	1.64	1.45	1.82	17.34		
No	44	2,067	1.88	1.64	2.12	15.36	2.47	0.116
IPT								
Yes	21	1,067	1.51	1.29	1.73	13.64		
No	13	453	1.80	1.47	2.13	10.66	1.99	0.159
Format								
CBT V IPT								
Individual	111	6,724	1.65	1.52	1.77	26.43	1.69	0.194
Group	30	2,890	1.33	1.15	1.50	15.03	0.42	0.519
CBT								
Individual	81	5,344	1.73	1.55	1.90	19.64		
Group	24	2,660	1.31	1.13	1.49	14.39	10.75	0.001

IPT								
Individual	30	1,380	1.57	1.39	1.74	17.74		
Group	6	230	1.55	0.86	2.24	4.39	0.01	0.942
				Setting				
CBT V IPT								
Outpatient	156	10,713	1.59	1.50	1.70	30.94	1.53	0.216
CBT								
Inpatient	6	422	1.36	0.87	1.84	5.44		
Outpatient	112	8,853	1.65	1.52	1.78	24.65	1.26	0.261
				Antidepressant Medication				
CBT V IPT								
Yes	57	6,389	1.47	1.31	1.64	17.55	1.32	0.251
No	84	3,680	1.68	1.55	1.81	24.78	4.35	0.037
CBT								
Yes	44	5,835	1.42	1.24	1.61	14.95		
No	57	2,586	1.82	1.63	2.01	19.13	8.74	0.003
IPT								
Yes	13	554	1.65	1.30	2.00	9.27		
No	27	1,094	1.54	1.35	1.73	15.89	0.34	0.562

Note. CI = confidence interval; *k*= number of samples; *N*= number of patients; *g*= Hedges' *g* effect size; *LL*= lower limit; *UL*= upper limit; *Zw*= within group heterogeneity; *Qb*= between group heterogeneity; *p*= significance value.

Table 3. Meta-regression mixed-effects anlayses of the following moderators: age, gender, employment, number of previous episodes and number of dropouts

					95% CI					
T	***k***	***N***	***b***	***Y***	***LL***	***UL***	***p***	Q_m	p_m	R^2
Age (mean in years)										
CBT & IPT	160	11,196	-0.009	1.95	-0.017	-0.000	.044	4.28	.118	.03
CBT	114	9,234	-0.011	2.05	-0.021	-0.002	.021	5.31	.021	.05
IPT	46	1,962	0.055	1.36	-0.011	0.023	.526	0.40	.526	.02
Gender (% male)										
CBT & IPT	163	11,251	0.003	1.53	-0.004	0.009	.404	0.91	.634	.00
Employment (% employed)										
CBT & IPT	44	3,317	0.003	1.37	-0.006	0.012	.501	0.54	.764	.00
Number of previous episodes (mean)										
CBT	20	830	0.003	1.77	-0.172	0.173	.997	0.00	.997	.00
Number of dropouts										
CBT & IPT	107	7,637	0.003	1.75	-0.007	0.001	.163	2.50	0.287	.00

Note. CI= confidence interval; T= therapy; *k*= number of samples; *N*= number of patients; *b*= predictor coefficient; *Y*= intercept; *LL*= lower limit; *UL*= upper limit; *p*= significance value of named predictor; Q_m= heterogeneity of the model; p_m= significance value of the model; R^2= coefficient of determination