Review

Mindfulness-based interventions for psychiatric disorders: A systematic review and meta-analysis

Simon B. Goldberg[a,b,c,*], Raymond P. Tucker[d], Preston A. Greene[a], Richard J. Davidson[b,e], Bruce E. Wampold[c,f], David J. Kearney[a], Tracy L. Simpson[a,g]

[a] VA Puget Sound Health Care System, Seattle, WA, USA
[b] Center for Healthy Minds, University of Wisconsin - Madison, Madison, WI, USA
[c] Department of Counseling Psychology, University of Wisconsin - Madison, Madison, WI, USA
[d] Department of Psychology, Louisiana State University, Baton Rouge, LA, USA
[e] Department of Psychology, University of Wisconsin - Madison, Madison, WI, USA
[f] Modum Bad Psychiatric Center, Vikersund, Norway
[g] Center for Excellence in Substance Abuse Treatment and Education, VA Puget Sound Health Care System, Seattle, WA, USA

HIGHLIGHTS

- We examined the relative efficacy of mindfulness-based interventions on clinical symptoms of psychiatric disorders.
- 142 randomized clinical trials were included (N = 12,005 participants). Control conditions were coded on a five-tier system.
- At post-treatment, mindfulness interventions were equivalent to evidence-based treatments and superior to other comparisons.
- At follow-up, mindfulness interventions were equivalent to minimal and evidence-based treatments and superior to others.
- The most consistent evidence for mindfulness-based interventions was seen for depression, pain, smoking, and addictions.

ARTICLE INFO

ABSTRACT

Despite widespread scientific and popular interest in mindfulness-based interventions, questions regarding the empirical status of these treatments remain. We sought to examine the efficacy of mindfulness-based interventions for clinical populations on disorder-specific symptoms. To address the question of relative efficacy, we coded the strength of the comparison group into five categories: no treatment, minimal treatment, non-specific active control, specific active control, and evidence-based treatment. A total of 142 non-overlapping samples and 12,005 participants were included. At post-treatment, mindfulness-based interventions were superior to no treatment ($d = 0.55$), minimal treatment ($d = 0.37$), non-specific active controls ($d = 0.35$), and specific active controls ($d = 0.23$). Mindfulness conditions did not differ from evidence-based treatments ($d = -0.004$). At follow-up, mindfulness-based interventions were superior to no treatment conditions ($d = 0.50$), non-specific active controls ($d = 0.52$), and specific active controls ($d = 0.29$). Mindfulness conditions did not differ from minimal treatment conditions ($d = 0.38$) and evidence-based treatments ($d = 0.09$). Effects on specific disorder subgroups showed the most consistent evidence in support of mindfulness for depression, pain conditions, smoking, and addictive disorders. Results support the notion that mindfulness-based interventions hold promise as evidence-based treatments.

1. Introduction

Mindfulness-based interventions have experienced a marked increase in scientific and popular interest in the past two decades. Recent commentaries have, however, raised questions regarding the evidence base for this family of therapies. Farias, Wikholm, and Delmonte (2016) voiced several concerns, particularly the use of non-active control conditions (i.e., waitlist controls) in randomized clinical trials (RCTs) of mindfulness therapies along with a lack of specificity regarding outcomes that these treatments may or may not impact. Others have questioned the degree to which selective reporting of results may introduce systematic bias into the literature, thereby overstating the

---

efficacy of mindfulness-based interventions (Coronado-Montoya et al., 2016).

One recent meta-analysis estimated the effects of meditation-based interventions (including mindfulness as well as other meditative techniques) compared to active control conditions that, analogous to placebos in pharmaceutical trials, provide non-specific treatment ingredients (e.g., expectancy; Goyal et al., 2014). While mindfulness meditation programs showed effects on anxiety, depression, and pain when compared with non-specific treatment controls, there was no evidence that these treatments were superior to specific active controls (i.e., other active treatments).

The current meta-analysis was intended to further interrogate the findings of Goyal et al. (2014). We conducted a comprehensive meta-analysis of RCTs examining the effects of mindfulness-based interventions on disorder-specific symptoms across psychiatric populations. Rather than restrict our sample to certain types of comparison conditions, we aimed to evaluate empirically the degree to which outcomes are influenced by the characteristics of the control group. A more nuanced comparison to type of control condition may provide clinicians important information regarding when a mindfulness intervention should be favored compared to other known interventions. While other comprehensive meta-analyses have suggested that mindfulness-based interventions can impact clinical outcomes (e.g., anxiety, depression; Khoury et al., 2013), and several meta-analyses have examined the evidence for specific psychiatric conditions (e.g., Attention Deficit and Hyperactivity Disorder [ADHD]; Cairncross & Miller, 2016), no published comprehensive meta-analytic review has examined effects on disorder-specific symptoms across psychiatric conditions. Our study sought to examine: (1) the degree to which mindfulness-based interventions compare with a variety of control conditions, including treatments with established efficacy (i.e., evidence-based treatments); (2) for which specific disorders mindfulness-based interventions appear most efficacious, and (3) potential sources of bias.

2. Method

2.1. Eligibility criteria

We included all RCTs of mindfulness-based interventions for adult patients with psychiatric diagnoses for which there are evidence-based treatments per the American Psychological Association's (APA, 2017) Division 12 (Society of Clinical Psychology; see Supplemental materials Table 1a). To be eligible, samples had to have either a formal diagnosis or elevated symptoms of a given disorder (i.e., above a given cut-off on a symptom inventory, e.g., score greater than five on the Pittsburgh Sleep Quality Index, score > 13 on the Beck Depression Inventory – II; Asl & Barahmand, 2014; Beck, Steer, & Brown, 1996; Black, O'Reilly, Olmstead, Breen, & Irwin, 2015; Buysse, Reynolds, Monk, Berman, & Kupfer, 1989). Samples receiving treatment within a facility focused on a specific disorder (e.g., substance abuse treatment) were included. Elevated stress levels alone were not considered to reflect a clinical condition.

To qualify, interventions had to have mindfulness meditation as a core component with home meditation practice as a treatment ingredient. While interventions combining mindfulness with other modalities (e.g., mindfulness and cognitive techniques as in Mindfulness-Based Cognitive Therapy [MBCT]; Segal, Williams, & Teasdale, 2002) were included, therapies emphasizing the attitudinal stance of mindfulness (rather than the formal practice of mindfulness meditation) were excluded (e.g., Acceptance and Commitment Therapy [ACT], Dialectical Behavior Therapy [DBT]; Hayes, Strosahl, & Wilson, 1999; Linehan, 1993). Other forms of meditation (e.g., mantram repetition) were excluded. Interventions had to be delivered in real time (i.e., not provided exclusively through video instruction or smartphone app) and had to include more than one session (to allow for home meditation practice). Studies were also excluded for the following reasons: (1) not

published in a peer-reviewed journal in English; (2) not a peer-reviewed article; (3) data unavailable to compute standardized effect sizes; (4) no disorder-specific (i.e., targeted) outcomes reported; (5) data redundant with other included studies; (6) no non-mindfulness-based intervention or condition included.

2.2. Information sources

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Standards (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009). We searched the four databases (PubMed, PsycInfo, Scopus, Web of Science) and a publically available comprehensive repository of mindfulness studies (Black, 2012). Citations from recent meta-analyses and systematic reviews were also reviewed. Citations were included from the first available date (i.e., 1966) until January 2nd, 2017.

2.3. Search

We used the search terms "mindfulness" and "random*". When a database allowed, we restricted our search to clinical trials.

2.4. Study selection

Titles and/or abstracts of potential studies were independently coded by the first author and a second co-author. Disagreements were discussed with the senior author until a consensus was reached.

2.5. Data collection process

Standardized spreadsheets were developed for coding both study-level and effect size-level data. Doctoral-level coders were trained by the first author through coding an initial sample of studies ($k = 10$) in order to achieve reliability. Data were extracted independently by the first author and a second co-author. Disagreements were discussed with the senior author. Inter-rater reliabilities were in the good to excellent range (i.e., Ks and ICCs > 0.60; Cicchetti, 1994). When sufficient data for computing standardized effect sizes were unavailable, study authors were contacted.

2.6. Data items

Along with data necessary for computing standardized effect sizes, the following data were extracted: (1) publication year; (2) disorder; (3) intent-to-treat (ITT) sample size; (4) whether an ITT analysis was reported; (5) whether a non-self-report measure was included; (6) sample demographics (mean age, percentage female, percentage non-Caucasian race, percentage with some college education); (7) country of origin; (8) standardized mindfulness intervention on which mindfulness condition was based; (9) whether treatment time was matched between mindfulness and control condition; (10) quality of the control condition. Quality of the control condition was assessed based on a five-tier system with non-overlapping categories. These included: (1) no treatment (in which the control condition received no intervention beyond that which was provided to the treatment condition); (2) minimal treatment (very brief or minimal intensity interventions, e.g., five- to 10-min individual counseling sessions for smoking cessation; Vidrine et al., 2016); (3) non-specific active control (active conditions in which no mechanism of change or clear rationale for treatment was provided, e.g., discussing air travel, shopping, and past residences; Helms & Ward, 2017); (4) specific active control (contained specific therapeutic mechanisms, has a theoretical/treatment rationale, e.g., Intensive Short-Term Dynamic Psychotherapy; Chavooshi, Mohammadkhani, & Dolatshahee, 2016; Wampold et al., 1997); (5) evidence-based treatment (EBT, e.g., cognitive-behavioral therapy for insomnia; Garland et al., 2014). Comparison treatments were coded as

EBTs if they were identified by APA Division 12[6] as an EBT for that particular disorder, or if they were promoted as a first-line treatment by a similarly relevant organization (e.g., smoking cessation treatment promoted by the American Lung Association, cognitive-behavioral therapy promoted by the National Institute on Drug Abuse).

2.7. Risk of bias in individual studies

Considerations for minimizing bias in individual studies were drawn from both Jadad's criteria as well as the GRADE system (Atkins et al., 2004; Jadad et al., 1996). Based on the GRADE recommendation to select relevant study characteristics to quantify (Agency for Healthcare Research and Quality, 2014) and based on the large number of potential study characteristics for assessing quality in psychotherapy trials, (e.g., $n = 185$ quality criteria; Liebherz, Schmidt, & Rabung, 2016), we restricted our analysis to randomized trials, employed intent-to-treat samples (when available), and coded the strength of the comparison condition (as described above), whether an ITT analysis was reported (e.g., using multiple imputation, last observation carried forward, or conservative assumptions regarding outcomes for participants who dropped out of the study [e.g., smoking relapse; Davis et al., 2014]), and whether a non-self-report outcome was included (e.g., biologically-confirmed abstinence, clinician-rated diagnostic status; Davis et al., 2014; Teasdale et al., 2000).

2.8. Summary measures

Our primary effect size measure was the standardized mean difference (Cohen's d). As done by Goyal et al. (2014), we first computed a pre-post effect size for both the mindfulness and non-mindfulness groups alone. This method has the advantage of accounting for potential baseline differences (i.e., it does not rely exclusively on between-group differences at post-test; Becker, 1988). We then calculated the relative difference in the pre-post effects (i.e., change scores) using standard methods (Becker, 1988), including controlling for a small known bias in d (Cooper, Hedges, & Valentine, 2009). Analyses were conducted using the R statistical software and the 'metafor' and 'MAd' packages (Del Re & Hoyt, 2010; Viechtbauer, 2010). Cohen's d was computed both from pre- to post-treatment (or time point closest to post-treatment) as well as from pre- to last available follow-up time point. Random effects models were used.

2.9. Synthesis of results

When available, effect sizes were computed using pre- and post-test means and standard deviations (SD). Other reported statistics (e.g., F, t, p, odds ratios) were used when appropriate based on standard meta-analytic methods (Cooper et al., 2009). Data were aggregated first within-studies (i.e., across disorder-specific outcomes within a given study) using the 'MAd' package and then between studies, based on the comparison of interest using restricted maximum likelihood random effects analyses. Summary statistics were computed in Cohen's d units along with 95% confidence intervals. Heterogeneity was systematically assessed using the I^2 (measuring the proportion of between-study heterogeneity) and the Q-statistic (assessing whether between-study heterogeneity exceeds that expected by chance alone).

To answer the question of the degree to which mindfulness-based interventions demonstrate relative efficacy with other comparison group types, summary results were first aggregated across studies employing a given comparison condition type (e.g., specific active control conditions). Although this involved pooling outcomes across a variety of disorders, we believe this analysis most directly examines the degree to which mindfulness-based interventions compare, on average, with various control group types, including other active therapies and evidence-based treatments. Then, in order to examine relative efficacy at the disorder level, studies that shared a given comparison type (e.g., no

treatment controls) and a given disorder (e.g., depression) were analyzed separately. In order to more efficiently and reliably summarize results, specific conditions with similar core features were collapsed (e.g., anxiety disorders, addictive disorders). Disorder categories were based on the Diagnostic and Statistical Manual of Mental Disorders 5th edition (American Psychiatric Association, 2013). We employed the recommended convention of requiring at least four studies within a subgroup for moderator and subgroup analyses (Fu et al., 2011).

Some studies included multiple control groups ($k = 22$, e.g., Bowen et al., 2014), comorbid diagnoses ($k = 5$, e.g., depression and pain; De Jong et al., 2016), or both ($k = 1$, e.g., Zautra et al., 2008). We attempted to code and analyze these studies in ways that allowed their data to be most fully characterized (this was deemed preferable to ignoring data from either multiple control groups or on comorbid disorders). Specifically, when multiple control groups were included, data from the mindfulness conditions were replicated to allow a representation of the unique comparison with each control group. For samples with comorbid conditions, separate effect sizes were included for each disorder. In order to assess potential bias introduced by this, sensitivity analyses were run excluding multiple comparison groups (only the most rigorous of the comparison groups was retained in these analyses) and excluding outcomes on comorbid conditions (only one of the two comorbid conditions was retained in these analyses).

2.10. Risk of bias across studies

We assessed publication bias by visually inspecting funnel plots for asymmetry within the comparison of interest and by re-estimating models using trim-and-fill methods that account for the asymmetric distribution of studies around an omnibus effect (Viechtbauer, 2010). In addition, we ran models within comparison condition assessing whether various features of study quality (i.e., based on Jadad and GRADE guidelines; Atkins et al., 2004; Jadad et al., 1996) were related to outcome. These features included for whether an ITT analysis was reported, whether non-self-report measures were included, and whether treatment time was matched between the mindfulness and the comparison conditions.

3. Results

3.1. Study selection

A total of 9067 citations were retrieved. After 3485 duplicates were removed, 5582 unique titles and/or abstracts were coded. Following the application of the exclusion criteria (see PRISMA flow diagram in Fig. 1), 171 studies were retained for analysis representing 142 independent samples, 172 unique comparisons (some studies included multiple comparison groups and comorbid samples), and 12,005 participants. Included studies were published between 2000 and 2016.

3.2. Study characteristics

The aggregate effect size (d) and other study characteristics for each study are shown in Supplemental materials. The sample was on average 43.63 years old, 64.38% female, with 61.27% having some post-secondary education. The largest percentage of trials was conducted in the United States (44.44%). The largest proportion of studies used no treatment comparison conditions (52.10%). The most commonly studied disorder was depression (30.41%; see Supplemental Materials Fig. 1a). The majority of studies included a follow-up assessment time point ($k = 79$, 55.63%). For studies with a follow-up assessment, the average follow-up length post-treatment was 6.43 months ($SD = 5.36$, range = one to 24).

```
┌─────────────────────────────┐
│ Studies retrieved from      │
│ electronic database searches│        ┌──────────────────────────┐
│ 4 792 AMRA                  │        │ 596 from hand searching  │
│ 228 PsycInfo                │        └──────────────────────────┘
│ 637 PubMed                  │
│ 1 275 Scopus                │
│ 1 539 Web of Science        │
└─────────────────────────────┘
```

3 485 duplicates removed

5 582 title-abstract reviews

Excluded at title-abstract review level

3 620 not randomized clinical trial
1 107 not relevant diagnosis
286 not mindfulness-based intervention
157 not adults
35 not in person
31 not published in English
9 no non-mindfulness intervention
31 not peer-reviewed article
3 no targeted outcomes

303 full text reviews

Excluded at full text review level

26 not randomized clinical trial
25 not relevant diagnosis
22 not mindfulness-based intervention
1 not in person
1 not published in English
1 no non-mindfulness intervention
2 not peer-reviewed article
26 no targeted outcomes
28 no new data provided

171 included articles

3.3. Risk of bias within studies

All included studies used randomized designs. A minority of comparisons (41.32%) matched treatment time between the mindfulness and control conditions. Approximately half of the studies reported at least one ITT analysis (54.86%) and included at least one non-self-report measure (48.61%).
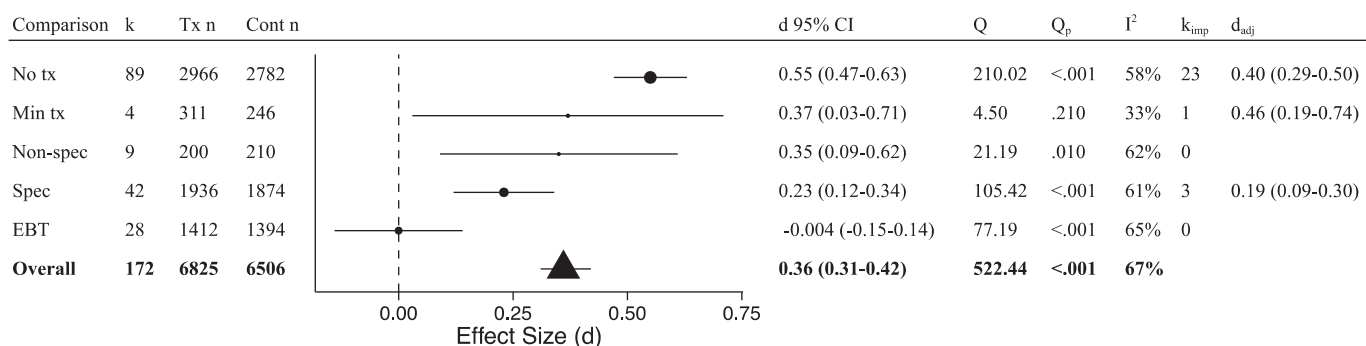
3.4. Results of individual studies

For each included study, treatment effects on disorder-specific outcomes and confidence intervals are reported in Supplemental Materials Table 2a. All included outcome measures for each study is listed in Supplemental materials Table 3a.

3.5. Synthesis of results

3.5.1. Effects at post-treatment

As expected, type of control condition was a significant moderator of effects at post-treatment ($Q[4] = 51.59$, $P < 0.001$; Fig. 2). Mindfulness-based interventions were shown to be superior to no treatment conditions ($k = 89$, $d = 0.55$ 95% CI 0.47 to 0.63), minimal treatment conditions ($k = 4$, $d = 0.37$ 95% CI 0.03 to 0.71), non-specific treatment conditions ($k = 9$, $d = 0.35$ 95% CI 0.09 to 0.62), and specific treatment conditions ($k = 42$, $d = 0.23$ 95% CI 0.12 to 0.34). Mindfulness-based interventions did not differ from EBTs ($k = 28$, $d = -0.004$ 95% CI -0.15 to 0.14). Within each comparison significant heterogeneity was detected, with the exception of minimal treatment comparisons.

Disorder type was next examined as a moderator of effects for

Comparison	k	Tx n	Cont n		d 95% CI	Q	Q_p	I^2	k_{imp}	d_{adj}
No tx	89	2966	2782		0.55 (0.47-0.63)	210.02	<.001	58%	23	0.40 (0.29-0.50)
Min tx	4	311	246		0.37 (0.03-0.71)	4.50	.210	33%	1	0.46 (0.19-0.74)
Non-spec	9	200	210		0.35 (0.09-0.62)	21.19	.010	62%	0	
Spec	42	1936	1874		0.23 (0.12-0.34)	105.42	<.001	61%	3	0.19 (0.09-0.30)
EBT	28	1412	1394		-0.004 (-0.15-0.14)	77.19	<.001	65%	0	
Overall	**172**	**6825**	**6506**		**0.36 (0.31-0.42)**	**522.44**	**<.001**	**67%**		

Effect Size (d): 0.00 0.25 0.50 0.75

Fig. 2. Post-treatment effects by comparison group type. k = number of disorder-specific comparisons; Tx n = mindfulness condition sample size; Cont n = comparison group sample size (note that total n is larger than the full sample size as some samples are represented in multiple comparisons); d = Cohen's d effect size; Q = Q-statistic; Q_p = p-value for Q-statistic; I^2 = heterogeneity; k_{imp} = number of imputed studies based on trim-and-fill analyses; d_{adj} = adjusted d based on trim-and-fill analyses; No tx = no treatment; Min tx = minimal treatment; Non-spec = non-specific active control condition; Spec = Specific active control condition; EBT = evidence-based treatment.

Comp	Diagnosis	k	Tx n	Cont n		d 95% CI	Q	Q_p	I^2	k_{imp}	d_{adj}
No tx	Addiction	5	80	69		0.35 (-0.06-0.76)	12.97	.010	69%	0	
No tx	Anxiety	8	250	222		0.89 (0.62-1.17)	36.49	<.001	81%	0	
No tx	Depression	30	1224	1145		0.59 (0.46-0.73)	44.82	.031	35%	9	0.45 (0.32-0.58)
No tx	Pain	24	782	752		0.45 (0.30-0.60)	42.88	.007	46%	6	0.31 (0.14-0.48)
No tx	Schizophrenia	7	228	228		0.50 (0.23-0.76)	11.04	.087	46%	0	
No tx	Weight/Eating	5	116	110		0.79 (0.44-1.15)	22.27	<.001	82%	0	
No tx	**Overall**	**79**	**2680**	**2526**		**0.56 (0.48-0.65)**	**187.59**	**<.001**	**58%**		
Spec	Addiction	7	482	418		0.27 (0.02-0.53)	7.28	.295	18%	0	
Spec	Anxiety	5	190	184		0.15 (-0.16-0.46)	4.69	.321	15%	0	
Spec	Depression	9	377	385		0.38 (0.12-0.64)	28.88	<.001	72%	0	
Spec	Pain	9	471	475		0.03 (-0.20-0.26)	38.04	<.001	79%	0	
Spec	Weight/Eating	5	221	216		0.08 (-0.24-0.39)	2.86	.582	0%	0	
Spec	**Overall**	**35**	**1741**	**1678**		**0.18 (0.06-0.30)**	**85.7**	**<.001**	**60%**		
EBT	Anxiety	5	174	188		-0.18 (-0.41-0.06)	6.4	.171	38%	1	-0.13 (-0.41-0.15)
EBT	Depression	10	528	536		-0.01 (-0.19-0.16)	8.85	.451	0%	1	0.01 (-0.16-0.19)
EBT	Smoking	4	293	294		0.42 (0.20-0.64)	3.36	.339	11%	2	0.32 (0.13-0.52)
EBT	**Overall**	**19**	**995**	**1018**		**0.07 (-0.05-0.19)**	**35.34**	**.009**	**49%**		

Effect Size (d): 0.0 0.5 1.0
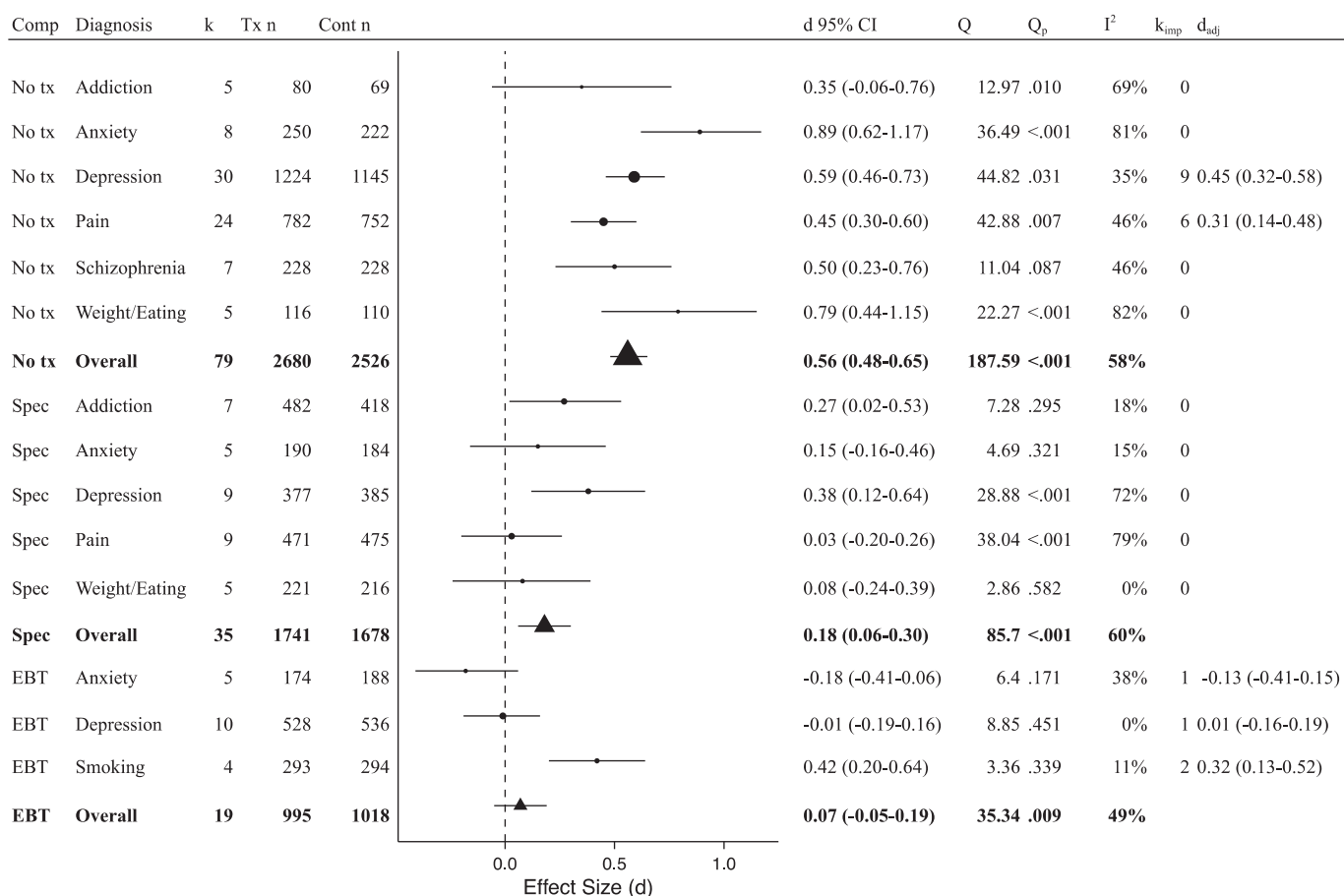
Fig. 3. Post-treatment effects on disorder-specific symptoms by comparison group and disorder. Comp = comparison group; k = number of disorder-specific comparisons; Tx n = mindfulness condition sample size; Cont n = comparison group sample size (note that total n is larger than the full sample size as some samples are represented in multiple comparisons); d = Cohen's d effect size; Q = Q-statistic; Q_p = p-value for Q-statistic; I^2 = heterogeneity; k_{imp} = number of imputed studies based on trim-and-fill analyses; d_{adj} = adjusted d based on trim-and-fill analyses; No tx = no treatment; Min tx = minimal treatment; Non-spec = non-specific active control condition; Spec = Specific active control condition; EBT = evidence-based treatment.

studies using the same comparison conditions. Disorders were included in this analysis if at least four trials employing a given comparison condition were available (Fu et al., 2011). For studies using a no treatment comparison, disorder was not a significant moderator (Q[5] = 10.86, p = 0.054). Mindfulness-based interventions showed superior effects on disorder-specific outcomes for anxiety, depression, pain, schizophrenia, and weight/eating-related disorders, with ds ranging from 0.45 to 0.89 (Fig. 3); addictions were the exception (d = 0.35 95% CI -0.06 to 0.76). Sufficient studies were not available for minimal

treatment or non-specific treatment comparison types. For specific active control conditions, disorder was not a significant moderator (Q[3] = 4.84, p = 0.305). Mindfulness-based interventions were superior to the comparison group for depression and addiction (ds = 0.27 to 0.38) and equivalent to the comparison group for anxiety, pain, and weight/eating (ds = 0.03 to 0.15). When compared with EBTs, disorder was a significant moderator (Q[2] = 14.51, p = 0.001). Mindfulness-based interventions were superior to EBTs for smoking (d = 0.42) and equivalent to EBTs for anxiety and depression (ds = − 0.01 to − 0.18).
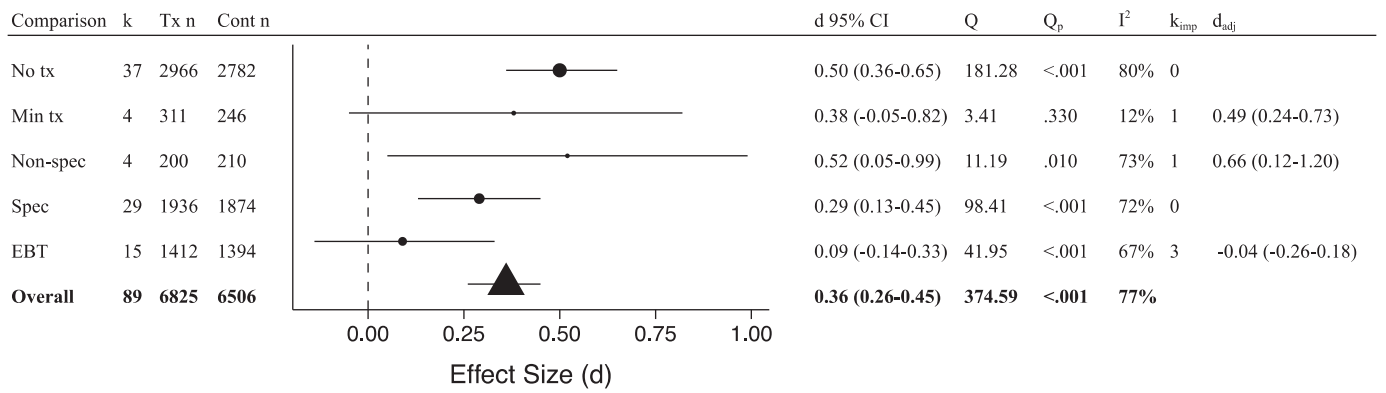
Comparison	k	Tx n	Cont n		d 95% CI	Q	Q_p	I^2	k_{imp}	d_{adj}
No tx	37	2966	2782		0.50 (0.36-0.65)	181.28	<.001	80%	0	
Min tx	4	311	246		0.38 (-0.05-0.82)	3.41	.330	12%	1	0.49 (0.24-0.73)
Non-spec	4	200	210		0.52 (0.05-0.99)	11.19	.010	73%	1	0.66 (0.12-1.20)
Spec	29	1936	1874		0.29 (0.13-0.45)	98.41	<.001	72%	0	
EBT	15	1412	1394		0.09 (-0.14-0.33)	41.95	<.001	67%	3	-0.04 (-0.26-0.18)
Overall	**89**	**6825**	**6506**		**0.36 (0.26-0.45)**	**374.59**	**<.001**	**77%**		

Effect Size (d)

Fig. 4. Follow-up effects by comparison type. k = number of disorder-specific comparisons; Tx n = mindfulness condition sample size; Cont n = comparison group sample size (note that total n is larger than the full sample size as some samples are represented in multiple comparisons); d = Cohen's d effect size; Q = Q-statistic; Q_p = p-value for Q-statistic; I^2 = heterogeneity; k_{imp} = number of imputed studies based on trim-and-fill analyses; d_{adj} = adjusted d based on trim-and-fill analyses; No tx = no treatment; Min tx = minimal treatment; Non-spec = non-specific active control condition; Spec = Specific active control condition; EBT = evidence-based treatment.

3.5.2. Effects at longest follow-up

At follow-up, type of control condition was a significant moderator of effects ($Q[4] = 9.85$, $p = 0.043$; Fig. 4). Mindfulness-based interventions were shown to be superior to no treatment conditions ($k = 37$, $d = 0.50$ 95% CI 0.36 to 0.65), non-specific treatment conditions ($k = 4$, $d = 0.52$ 95% CI 0.05 to 0.99), and specific active controls ($k = 29$, $d = 0.29$ 95% CI 0.13 to 0.45). Mindfulness-based interventions did not differ statistically from minimal treatment conditions ($k = 4$, $d = 0.38$ 95% CI -0.05 to 0.82) and EBTs ($k = 15$, $d = 0.09$ 95% CI -0.14 to 0.33). Within each comparison, significant heterogeneity was detected (Fig. 4), with the exception of the minimal treatment comparisons.

Disorder type was again examined as a moderator of effects for studies using the same comparison conditions. For studies using a no treatment comparison, disorder was a significant moderator ($Q[2] = 6.46$, $p = 0.040$). Mindfulness-based interventions showed superior effects for depression, pain, and schizophrenia, with ds ranging from

0.48 to 1.18 (Fig. 5). Sufficient studies were not available for minimal treatment or non-specific treatment comparison types. For specific active control conditions, disorder was not a significant moderator ($Q[3] = 1.22$, $p = 0.748$). Mindfulness-based interventions were superior to the comparison group for depression ($d = 0.35$) and equivalent to the comparison group for addictions, pain, and weight/eating ($ds = 0.18$ to 0.38). Mindfulness-based interventions were equivalent to EBTs for depression ($d = 0.04$).

3.6. Sensitivity analyses

A series of sensitivity analyses examined the potential impact of duplicating data from portions of the sample in order to allow for multiple comparison groups and/or multiple disorder-specific outcomes to be fully represented. Results from these models are reported in Supplementary Materials and are summarized here. When excluding multiple comparison groups at post-treatment, insufficient studies were
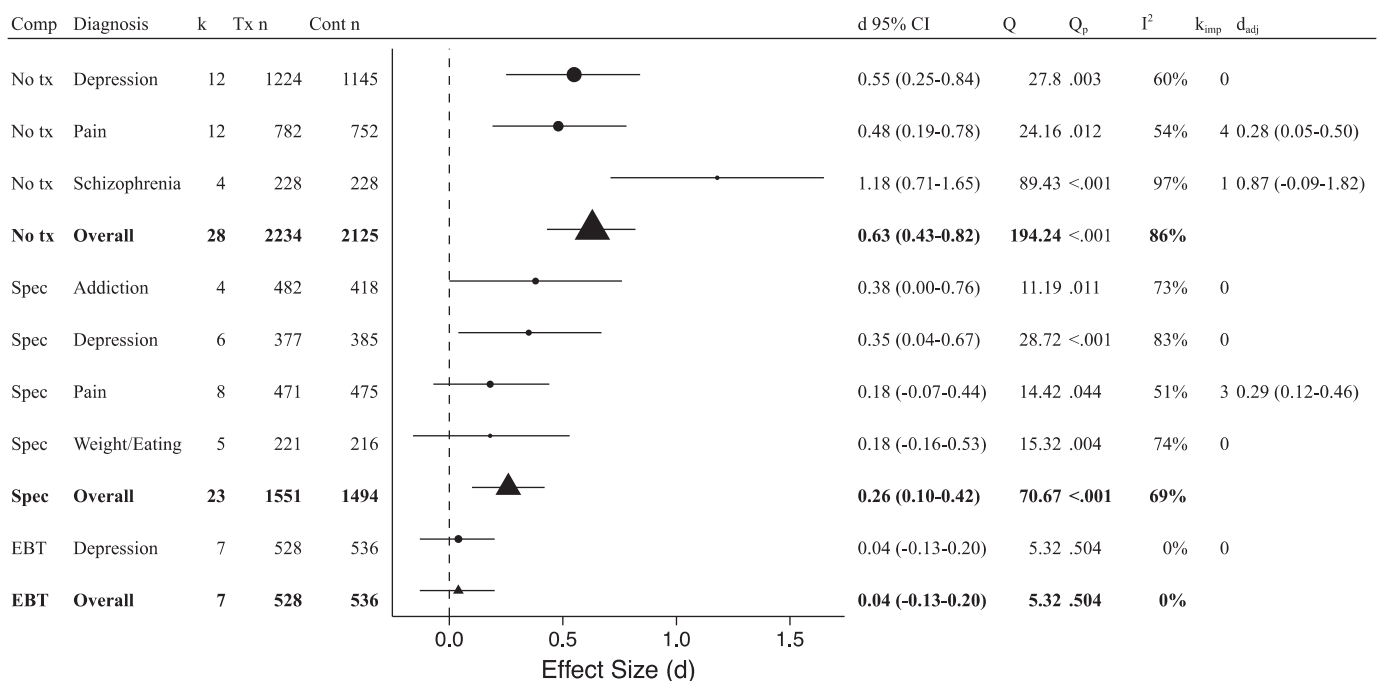
Comp	Diagnosis	k	Tx n	Cont n		d 95% CI	Q	Q_p	I^2	k_{imp}	d_{adj}
No tx	Depression	12	1224	1145		0.55 (0.25-0.84)	27.8	.003	60%	0	
No tx	Pain	12	782	752		0.48 (0.19-0.78)	24.16	.012	54%	4	0.28 (0.05-0.50)
No tx	Schizophrenia	4	228	228		1.18 (0.71-1.65)	89.43	<.001	97%	1	0.87 (-0.09-1.82)
No tx	**Overall**	**28**	**2234**	**2125**		**0.63 (0.43-0.82)**	**194.24**	**<.001**	**86%**		
Spec	Addiction	4	482	418		0.38 (0.00-0.76)	11.19	.011	73%	0	
Spec	Depression	6	377	385		0.35 (0.04-0.67)	28.72	<.001	83%	0	
Spec	Pain	8	471	475		0.18 (-0.07-0.44)	14.42	.044	51%	3	0.29 (0.12-0.46)
Spec	Weight/Eating	5	221	216		0.18 (-0.16-0.53)	15.32	.004	74%	0	
Spec	**Overall**	**23**	**1551**	**1494**		**0.26 (0.10-0.42)**	**70.67**	**<.001**	**69%**		
EBT	Depression	7	528	536		0.04 (-0.13-0.20)	5.32	.504	0%	0	
EBT	**Overall**	**7**	**528**	**536**		**0.04 (-0.13-0.20)**	**5.32**	**.504**	**0%**		

Effect Size (d)

Fig. 5. Follow-up effects by comparison type and disorder. Comp = comparison group; k = number of disorder-specific comparisons; Tx n = mindfulness condition sample size; Cont n = comparison group sample size (note that total n is larger than the full sample size as some samples are represented in multiple comparisons); d = Cohen's d effect size; Q = Q-statistic; Q_p = p-value for Q-statistic; I^2 = heterogeneity; k_{imp} = number of imputed studies based on trim-and-fill analyses; d_{adj} = adjusted d based on trim-and-fill analyses; No tx = no treatment; Min tx = minimal treatment; Non-spec = non-specific active control condition; Spec = Specific active control condition; EBT = evidence-based treatment.

available to estimate minimal treatment comparisons or effects on weight/eating for no treatment comparisons. All significance tests remained unchanged, with the exception of effects on addiction for specific active control comparisons, which became non-significant, although the effect size increased slightly ($ds = 0.27$ 95% CI 0.02 to 0.53 and 0.29 95% CI -0.04 to 0.62, for multiple comparisons included and excluded, respectively). When excluding multiple groups at follow-up, insufficient studies were available to estimate minimal treatment comparisons and non-specific treatment comparisons. Insufficient studies were available for estimating effects on schizophrenia for no treatment comparisons and effects on addiction for specific active controls. All remaining significance tests were unchanged.

For models excluding multiple disorders, the significance tests for all analyses remained unchanged with the exception of studies of depression using specific active controls at follow-up which became non-significant ($ds = 0.35$ 95% CI 0.04 to 0.67 and 0.15 95% CI -0.13 to 0.44, for multiple disorders included and excluded respectively).

3.7. Risk of bias across studies

Asymmetric funnel plots suggested evidence for publication bias for several analyses (see Supplemental materials Figs. 2a to 7a), with trim-and-fill analyses resulting in modified effect size estimates. Of note, the statistical significance of all estimates remained unchanged, with the exception of no treatment comparisons for schizophrenia at longest follow-up, which was no longer significant. Neither reporting ITT analyses, including non-self-report outcomes, nor matching treatment time between mindfulness and comparison conditions predicted outcomes when examined within comparison type significantly moderated effects at post-treatment or follow-up (all $ps > 0.05$, see Supplemental Materials Table 7a), with one exception: studies using non-specific active controls and reporting objective outcomes had significantly lower effects at follow-up ($d = 0.25$) relative to those not using objective outcomes ($d = 1.34$, $Q[1] = 10.08$, $p = 0.002$).

4. Discussion

Several conclusions can be drawn from these findings. At the most basic level, our results suggest that there is an empirical basis for mindfulness-based therapies. Mindfulness treatments were shown, in general, to be of similar potency with first-line psychological (and psychiatric) interventions when compared directly and superior to other active comparison conditions (as well as waitlist control conditions), with relatively little variation across disorders. These effects were generally robust to accounting for publication bias, study quality features, and in sensitivity analyses that restricted our sample to one comparison per study. This finding supports continued research exploring the clinical application of mindfulness therapies and provides a basis for consideration of these treatments by medical providers.

The promising effects demonstrated on psychiatric symptoms in the included studies are consistent with several other symptom- or disorder-specific meta-analytic reviews (e.g., Cairncross & Miller, 2016; Khoury et al., 2013; Khoury, Lecomte, Gaudiano, & Paquin, 2013; Piet & Hougaard, 2011) as well as with a comprehensive review in child and adolescent samples (Zoogman, Goldberg, Hoyt, & Miller, 2015). The magnitude of the effect sizes detected in the current study (e.g., $d = 0.55$ for mindfulness versus no treatment comparison conditions at post-treatment) suggests that mindfulness-based interventions are, on average, associated with moderate drops in psychiatric symptoms (based on Cohen's, 1988 guidelines). Interestingly, our findings diverge from those of Goyal et al. (2014) who reported no differences between mindfulness conditions and specific active control conditions. This discrepancy may be due to the current meta-analysis including only disorder-specific symptoms and increased statistical power to detect difference through including a larger number of trials and examining effects at the level of comparison condition type (i.e., not only

disaggregated by disorder).

This is the first comprehensive meta-analysis to examine effects of mindfulness-based interventions on symptoms specific to clinical disorders. In addition, we have attempted to grade the strength of comparison conditions to rigorously examine the relative efficacy of mindfulness-based interventions compared not only to no treatment but also to other active treatments that may be recommended. We believe that this method addresses the primary question facing clinicians who may be themselves providing or considering recommending mindfulness-based therapies: How do mindfulness interventions compare with other evidence-based treatments?

Our sample included a wide range of psychiatric conditions with behavioral therapies known to be efficacious. Overall, mindfulness therapies were superior to no treatment, minimal treatment (at post-treatment), non-specific active controls (i.e., psychological placebo groups), and specific active controls (i.e., other psychological treatments). Further, mindfulness-based interventions were on average not different from first-line, evidence-based therapies such as cognitive behavioral therapy and antidepressant medications.

Subsequent analyses examined the relative performance of mindfulness-based interventions within categories of clinical conditions at post-treatment (e.g., anxiety disorders, addictive disorders). As analyses were restricted to comparisons that included at least four RCTs, only a subset of conditions could be assessed. The clearest evidence was found regarding the use of mindfulness for depression. Mindfulness was found to be superior to no treatment, other active therapies, and equivalent to EBTs. For pain and weight/eating, mindfulness performed on par with other active therapies and was superior to no treatment controls. For schizophrenia, mindfulness outperformed no treatment control conditions. For anxiety, mindfulness outperformed no treatment control conditions and was equivalent to other active therapies, including EBTs. For smoking, mindfulness outperformed EBTs. Effects on addictions varied. Mindfulness was equivalent to no treatment controls although superior to other active therapies. This apparently contradictory finding is likely due to the small number of studies examining addictive disorders using no treatment control groups ($k = 5$) along with the small sample size included in this particular group of studies (*mean n* = 29.8) which yielded a wide confidence interval for this effect size estimate. Examination of effect sizes at post-treatment shows the expectedly larger effect on addictions for no treatment comparisons ($d = 0.35$) than for specific active control conditions ($d = 0.27$), despite the contrasting significance tests.

At follow-up, results were similar, although not identical. In these analyses, mindfulness was no different than minimal treatment controls, again perhaps due to the small number of studies included in this group. Mindfulness was again superior to no treatment and specific active control comparisons, and equivalent to EBTs. For specific disorders, mindfulness outperformed no treatment comparisons for depression, pain, and schizophrenia (although not when accounting for publication bias), and was equivalent or superior to other active treatments for addictions, depression, pain, and weight/eating. When compared with EBTs for depression, mindfulness was equivalent.

It remains difficult, however, to make firm recommendations based on the literature regarding for which particular disorders these therapies hold most promise (Farias et al., 2016). This is due to the heterogeneity in effects across disorders, the uneven distribution of studies across disorders, the relative scarcity of direct comparisons between mindfulness-based therapies and other first-line treatments (a study design feature that may not be improving despite repeated concerns voiced in the literature; Goldberg et al., (2017), and evidence of publication bias. Based on our findings, it appears that the strongest recommendation can be made for mindfulness treatments for depression with evidence also supporting the use of mindfulness for treating pain conditions, smoking, and addictive disorders.

The uneven distribution of studies across disorders and comparison types is a primary limitation of the current study. As is always the case

with meta-analyses, we were limited by the published literature (and given the scope of the project focused exclusively on studies published in peer-reviewed journals, i.e., did not include unpublished studies or dissertations), and therefore were unable to make firm conclusions regarding disorder groups that have received less research attention. In addition, for the purposes of generating reliable effect size estimates, related outcomes and related disorders were collapsed (e.g., pain intensity and pain functionality, obesity and eating disorders) which limited our ability to detect specific effects at the outcome and disorder level.

Several key questions for future research suggest themselves. One concerns the impact of practice duration. Insufficient data were available to include this in the meta-analysis but other basic research clearly indicates the importance of practice duration on basic biological measures (Wielgosz, Schuyler, Lutz, & Davidson, 2016). A second critical question is which individuals may be most benefited by mindfulness interventions? Are there certain individual difference characteristics that predict the magnitude of change with mindfulness interventions (Mascaro, Rilling, Negi, & Raison, 2013)? Again basic research underscores the importance of such individual differences. Collectively our findings underscore the potential promise of mindfulness-based interventions for psychiatric disorders.

Role of funding source

Contributors

SG drafted the protocol with BW, RD, and TS. SG, RT, and PG assessed the eligibility of the studies for inclusion, extracted data, and assessed risk of bias. SG developed the statistical code and did the analyses. All authors contributed to the interpretation of the findings. SG drafted the manuscript, to which all authors contributed. RD and SG obtained the funding.

Conflict of interests

RD is the founder, president, and serves on the board of directors for the non-profit organization, Healthy Minds Innovations, Inc. In addition, RD serves on the board of directors for the Mind and Life Institute.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2017.10.011.

References

Agency for Healthcare Research and Quality (2014). *Methods guide for effectiveness and comparative effectiveness reviews.* Rockville, MD: Agency for Healthcare Research and Quality.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

American Psychological Association (2017). Research-supported psychological treatments (Retrieved from: 2 January) https://www.div12.org/psychological-treatments/.

Asl, N. H., & Barahmand, U. (2014). Effectiveness of mindfulness-based cognitive therapy for co-morbid depression in drug-dependent males. *Archives of Psychiatric Nursing, 28*(5), 314–318.

Atkins, D., Eccles, M., Flottorp, S., Guyatt, G. H., Henry, D., Hill, S., ... The GRADE Working Group (2004). Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches The GRADE Working Group. *BMC Health Services Research, 4*(38), http://dx.doi.org/10.1186/1472-6963-4-38.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck depression inventory manual* (2nd ed.). San Antonio, TX: Psychological Corporation.

Becker, B. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology, 41*, 257–278.

Black, D. S. (2012). Mindfulness research guide: A new paradigm for managing empirical health information. *Mindfulness, 1*(3), 174–176. http://dx.doi.org/10.1007/s12671-010-0019-0.

Black, D. S., O'Reilly, G. A., Olmstead, R., Breen, E. C., & Irwin, M. R. (2015). Mindfulness meditation and improvement in sleep quality and daytime impairment among older adults with sleep disturbances: A randomized clinical trial. *JAMA Internal Medicine, 175*(4), 494–501. http://dx.doi.org/10.1001/jamainternmed.2014.8081.

Bowen, S., Witkiewitz, K., Clifasefi, S. L., Grow, J., Chawla, N., Hsu, S. H., ... Larimer, M. E. (2014). Relative efficacy of mindfulness-based relapse prevention, standard relapse prevention, and treatment as usual for substance use disorders: A randomized clinical trial. *JAMA Psychiatry, 71*(5), 547–556.

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193–213.

Cairncross, M., & Miller, C. J. (2016). The effectiveness of mindfulness-based therapies for ADHD: A meta-analytic review. *Journal of Attention Disorders.* (epub ahead of print) https://doi.org/10.1177/1087054715625301.

Chavooshi, B., Mohammadkhani, P., & Dolatshahee, B. (2016). Efficacy of intensive short-term dynamic psychotherapy for medically unexplained pain: A pilot three-armed randomized controlled trial comparison with mindfulness-based stress reduction. *Psychotherapy and Psychosomatics, 85*(2), 123–125.

Cicchetti, D. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284–290.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cooper, H. M., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russell Sage Foundation.

Coronado-Montoya, S., Levis, A. W., Kwakkenbos, L., Steele, R. J., Turner, E. H., & Thombs, B. D. (2016). Reporting of positive results in randomized controlled trials of mindfulness-based mental health interventions. *PLoS One, 11*(4), e0153220. http://dx.doi.org/10.1371/journal.pone.0153220.

Davis, J. M., Goldberg, S. B., Anderson, M. C., Manley, A. R., Smith, S. S., & Baker, T. B. (2014). Randomized trial on mindfulness training for smokers targeted to a disadvantaged population. *Substance Use and Misuse, 49*, 571–585. http://dx.doi.org/10.3109/10826084.2013.770025.

Del Re, A. C., & Hoyt, W. T. (2010). *MAd: Meta-analysis with mean differences* (R package version 0)8. http://CRAN.R-project.org/package=MAd.

Farias, M., Wikholm, C., & Delmonte, R. (2016). What is mindfulness-based therapy good for? *The Lancet Psychiatry, 3*, 1012–1013.

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J., ... Santaguida, P. (2011). Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *Journal of Clinical Epidemiology, 64*(11), 1187–1197. http://dx.doi.org/10.1016/j.jclinepi.2010.08.010.

Garland, S. N., Carlson, L. E., Stephens, A. J., Antle, M. C., Samuels, C., & Campbell, T. S. (2014). Mindfulness-based stress reduction compared with cognitive behavioral therapy for the treatment of insomnia comorbid with cancer: A randomized, partially blinded, non inferiority trial. *Journal of Clinical Oncology, 32*(5), 449–457.

Goldberg, S. B., Tucker, R. P., Greene, P. A., Simpson, T. L., Kearney, D. J., & Davidson, R. J. (2017). Is mindfulness research methodology improving over time? A systematic review. *PLoS One, 12*(10), e0187298. http://dx.doi.org/10.1371/journal.pone.0187298.

Goyal, M., Singh, S., Sibinga, E. M., Gould, N. F., Rowland-Seymour, A., Sharma, R., ... Haythornthwaite, J. A. (2014). Meditation programs for psychological stress and well-being: A systematic review and meta-analysis. *JAMA Internal Medicine, 174*(3), 357–368. http://dx.doi.org/10.1001/jamainternmed.2013.13018.

Hayes, S. C., Strosahl, K., & Wilson, K. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change.* New York: Guilford Press.

Helmes, E., & Ward, B. G. (2017). Mindfulness-based cognitive therapy for anxiety symptoms in older adults in residential care. *Aging & Mental Health, 21*(3), 272–278.

Jadad, A. R., Moore, A., Carroll, D., Jenkinson, C., Reynolds, D. J. M., Gavaghan, D. J., et al. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials, 17*, 1–12.

de Jong, M., Lazar, S. W., Hug, K., Mehling, W. E., Hölzel, B. K., Sack, A. T., ... Gard, T. (2016). Effects of mindfulness-based cognitive therapy on body awareness in patients with chronic pain and comorbid depression. *Frontiers in Psychology, 7*, 1–13.

Khoury, B., Lecomte, T., Fortin, G., Masse, M., Therien, P., Bouchard, V., ... Hofmann, S. G. (2013). Mindfulness-based therapy: A comprehensive meta-analysis. *Clinical Psychology Review, 33*, 763–771. http://dx.doi.org/10.1016/j.cpr.2013.05.005.

Khoury, B., Lecomte, T., Gaudiano, B. A., & Paquin, K. (2013). Mindfulness intervention for psychosis: A meta-analysis. *Schizophrenia Research, 150*, 176–184. http://dx.doi.org/10.1016/j.schres.2013.07.055.

Liebherz, S., Schmidt, N., & Rabung, S. (2016). How to assess the quality of psychotherapy outcome studies: A systematic review of quality assessment criteria. *Psychotherapy Research, 26*(5), 573–589. http://dx.doi.org/10.1080/10503307.2015.1044763.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.*

New York: Guilford Press.

Mascaro, J. S., Rilling, J. K., Negi, L. T., & Raison, C. L. (2013). Pre-existing brain function predicts subsequent practice of mindfulness and compassion meditation. *NeuroImage, 69*, 35–42. http://dx.doi.org/10.1016/j.neuroimage.2012.12.021.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264–269.

Piet, J., & Hougaard, E. (2011). The effect of mindfulness-based cognitive therapy for prevention of relapse in recurrent major depressive disorder: A systematic review and meta-analysis. *Clinical Psychology Review, 31*(6), 1032–1040.

Segal, Z., Williams, J. W., & Teasdale, J. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to preventing relapse.* New York: Guilford Press.

Teasdale, J., Segal, Z., Williams, J., Ridgeway, V., Soulsby, J., & Lau, M. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology, 68*(4), 615–623.

Vidrine, J. I., Spears, C. A., Heppner, W. L., Reitzel, L. R., Marcus, M. T., Cinciripini, P. M., ... Tindle, H. A. (2016). Efficacy of mindfulness-based addiction treatment (MBAT) for smoking cessation and lapse recovery: A randomized clinical trial. *Journal of*

Consulting and Clinical Psychology, 84(9), 824–838. http://dx.doi.org/10.1037/ccp0000117.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–49.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes.". *Psychological Bulletin, 122*, 203–215.

Wielgosz, J., Schuyler, B. S., Lutz, A., & Davidson, R. J. (2016). Long-term mindfulness training is associated with reliable differences in resting respiration rate. *Scientific Reports, 6*(January), 27533. http://dx.doi.org/10.1038/srep27533.

Zautra, A. J., Davis, M. C., Reich, J. W., Nicassario, P., Tennen, H., Finan, P., ... Irwin, M. R. (2008). Comparison of cognitive behavioral and mindfulness meditation interventions on adaptation to rheumatoid arthritis for patients with and without history of recurrent depression. *Journal of Consulting and Clinical Psychology, 76*(3), 408–421.

Zoogman, S., Goldberg, S. B., Hoyt, W. T., & Miller, L. (2015). Mindfulness interventions with youth: A meta-analysis. *Mindfulness, 6*, 290–302. http://dx.doi.org/10.1007/s12671-013-0260-4.