Personality Disorders: Theory, Research, and Treatment

A Meta-Analysis of Psychodynamic Treatments for Borderline and Cluster C Personality Disorders

John R. Keefe, Shelley F. McMain, Kevin S. McCarthy, Sigal Zilcha-Mano, Ulrike Dinger, Zeynep Sahin, Kathryn Graham, and Jacques P. Barber

Online First Publication, December 5, 2019. http://dx.doi.org/10.1037/per0000382

American Psychological Association

Personality Disorders: Theory, Research, and Treatment

ISSN: 1949-2715

http://dx.doi.org/10.1037/per0000382

A Meta-Analysis of Psychodynamic Treatments for Borderline and Cluster C Personality Disorders

John R. Keefe
Weill Medical College of Cornell University

Shelley F. McMain
University of Toronto and Centre for Addiction and Mental Health, Toronto, Ontario, Canada

Kevin S. McCarthy
Chestnut Hill College

Sigal Zilcha-Mano
University of Haifa

Ulrike Dinger
Heidelberg University Hospital, Heidelberg, Germany

Zeynep Sahin, Kathryn Graham, and Jacques P. Barber
Adelphi University

Personality disorders (PD) carry high psychosocial dysfunction and are associated with treatment resistance in nonspecialized care. Psychodynamic therapies (PDT) are often used to treat PDs, but there has never been a systematic meta-analysis of PDT trials for PD. To evaluate the evidence base for PDTs for PDs across multiple outcome domain, a systematic search for PDT for PD trials was conducted through PubMed and PsycINFO. Sixteen trials were identified, comprising 19 dynamic, 8 active, and 9 control groups predominantly reflecting treatment of borderline and mixed Cluster C PDs, and a random effects meta-analysis was undertaken. PDTs were superior to controls in improving core PD symptoms ($g = -0.63$; 95% confidence interval [CI; −0.87, −0.41]), suicidality ($g = -0.79$, $p = .02$; 95% CI [−1.38, −0.20]), general psychiatric symptoms ($g = -0.47$; 95% CI [−0.69, −0.25]), and functioning ($g = -0.66$; 95% CI [−1.01, −0.32]), but not for interpersonal problems due to heterogeneity ($g = -1.25$; 95% CI [−3.22, 0.71]). Outcomes for PDTs were not different from other active treatments in core PD ($g = 0.05$; 95% CI [−0.25, 0.35]) or other symptoms. This pattern continued into posttreatment follow-up (average 14 months). Study quality was generally rated as adequate and was unrelated to outcomes. Compared with other treatments, PDTs do not have different acute effects and are superior to controls, although only trials treating BPD employed active controls and non-BPD trials were of lower quality. Underresearched areas include narcissistic PD, specific Cluster C disorders, and personality pathology as a continuous construct.

Keywords: meta-analysis, personality disorder, psychodynamic therapy, outcome research

Supplemental materials: http://dx.doi.org/10.1037/per0000382.supp

Personality disorders (PD) are prevalent mental illnesses (6.1–9.1%; Huang et al., 2009; Lenzenweger, Lane, Loranger, & Kessler, 2007) carrying high psychosocial burden (Ansell, Sanislow, McGlashan, & Grilo, 2007; Skodol et al., 2005). In the treatment of acute Axis I or symptom disorders (e.g., major depression), comorbid PDs predict lower remission rates for the symptom disorder (Ansell et al., 2011; Newton-Howes et al., 2014), greater resistance to the work of therapy (Zickgraf et al.,

John R. Keefe, Department of Psychiatry, Weill Medical College of Cornell University; Shelley F. McMain, Department of Psychiatry, University of Toronto, and Borderline Personality Clinic, Centre for Addiction and Mental Health, Toronto, Ontario, Canada; Kevin S. McCarthy, Department of Psychology, Chestnut Hill College; Sigal Zilcha-Mano, Department of Psychology, University of Haifa; Ulrike Dinger, Center for Psychosocial Medicine, Heidelberg University Hospital, Heidelberg, Germany; Zeynep Sahin, Kathryn Graham, and Jacques P. Barber, Gordon F. Derner School of Psychology, Adelphi University.

Shelley F. McMain is the primary investigator for one of the clinical trials included in the meta-analysis (McMain et al., 2009, 2012). Shelley F. McMain did not participate in any coding of study effects, quality, or moderators for her trial. PROSPERO International Prospective Register of Systematic Reviews registration number of the article is CRD42018077748.

John R. Keefe is supported by National Institutes of Health/National Center for Advancing Translational Sciences Grant TL1-TR-002386 through the Clinical and Translational Science Center at Weill Medical College of Cornell University.

Correspondence concerning this article should be addressed to John R. Keefe, Department of Psychiatry, Weill Medical College of Cornell University, 425 East 61st Street, Penthouse Floor, New York, NY 10065. E-mail: jack.keefe@gmail.com

2016), and greater likelihood of relapse and poorer functioning at follow-up (Grilo et al., 2010; Markowitz et al., 2007). For patients with symptom and personality disorders who are treated with non-PD focused treatments, remission rates for PDs are often lower than that for corresponding symptom disorders (Keefe, Milrod, Gallop, Barber, & Chambless, 2018; Keefe, Webb, & DeRubeis, 2016). Difficulties treating PD with common mood and anxiety disorder therapies bespeak the need for PD-focused treatments (cf. Magnavita, Levy, Critchfield, & Lebow, 2010).

Psychodynamic therapies (PDTs) are one of two families of therapies that have been tested multiple times in clinical trials for PDs, in addition to cognitive–behavioral treatments like dialectical behavior therapy (Linehan, 1993) and schema-focused therapy (Young, Klosko, & Weishaar, 2003). PDTs for PDs generally have the aim of helping patients ameliorate personality functioning and pathological ways of relating to self and other, including improving attachment, reflective functioning/mentalization, personality organization, and use of defense mechanisms (Keefe & DeRubeis, 2019). PDTs for PDs focus on particular mental contents (e.g., self-and-other representations), experiences (e.g., strong affects), and dynamic processes surrounding those features that maintain a symptom or personality constellation (e.g., emphasis of aggressive self-representations to defend against vulnerable self-representations; Barber, Muran, McCarthy, & Keefe, 2013). Psychodynamic interventions for PDs could (nonexhaustively) include interpretation of self-and-other representations and strong associated affects as they emerge outside the therapy and within a session (transference-focused psychotherapy; Yeomans, Clarkin, & Kernberg, 2015); exposure to defended-against affects (affect-phobia therapy; McCullough et al., 2003); modeling a mentalizing stance encouraging curiosity and meaning-making concerning self and other (mentalization-based treatment; Bateman & Fonagy, 2016); or reinforcing healthy aspects of personality and supporting reality-testing (good psychiatric management; Gunderson & Links, 2014). Increases in insight into psychological dynamics and defense (Johansson et al., 2010; Kallestad et al., 2010), in adaptive defense use (Johansen, Krebs, Svartberg, Stiles, & Holen, 2011; Perry & Bond, 2012), and tolerance of affects (Høglend & Hagtvet, 2019) have been found to potentially mediate or predict subsequent improvements in symptoms and functioning in PDTs for PD.

To date, there has not been a comprehensive meta-analysis of PDTs for all PDs using contemporary meta-analytic methods. An early effort to address PDT for PD outcomes estimated uncontrolled effect sizes between PDT and cognitive–behavioral trials for PD (Leichsenring & Leibing, 2003), finding that PDT trials typically reported higher uncontrolled effects. However, uncontrolled effect sizes are difficult to compare meaningfully between studies due to study-level differences, baseline variability in severity highly influencing the size of the calculated effect, and need for the (typically unavailable) prepost score correlation to properly calculate an effect size (Cuijpers, Weitz, Cristea, & Twisk, 2017; Keefe, 2015). Moreover, many included studies did not use a reliable measure for diagnosing PD, and the meta-analysis included studies of PD secondary to another psychiatric disorder that was the primary focus of treatment. A more recent meta-analysis of treatments for BPD concluded that dialectical behavioral therapy (DBT) and PDTs had the strongest available evidence for efficacy using between-groups, controlled effect sizes (Cristea et al., 2017). However, head-to-head comparisons between active treatments were not included in this meta-analysis. We aimed to systematically assess the evidence base of PDT for treatment of a primary PD (including BPD but also other PDs), as tested in randomized controlled trials (RCTs) comparing PDT to nondynamic, *bona fide* active treatments and to control groups. We sought to examine effects in specific symptom domains, including core PD symptoms, suicidality, general psychiatric symptoms (e.g., depression, anxiety), interpersonal problems, and functioning.

Method

Study Search

This protocol was preregistered through the PROSPERO International Prospective Register of Systematic Reviews (ID CRD42018077748).[1] Two study authors (John R. Keefe and Kathryn Graham) performed the study search. The primary study search was through the PubMed and PsycINFO databases. The search terms were as follows: ("personality disorder" OR "personality disorders" OR "BPD" OR "PD") AND ("psychodynamic" OR "psychoanalytic" or "dynamic" OR "transference" OR "mentalization") AND ("trial" OR "RCT" OR "randomized"). Furthermore, we consulted the Lilliengren list (Lilliengren, 2018), a regularly updated compendium of psychodynamic clinical trials, to check full coverage of relevant studies, in addition to past meta-analyses and reviews involving PDT for PD. Studies were included by consensus of both study searchers.

For the purpose of our study, psychodynamic therapy was defined as an umbrella category for therapies arrayed on a supportive-expressive continuum (Luborsky, 1984; see online supplemental materials). Inclusion criteria for the meta-analysis included studies that (a) were published after 1970; (b) were in the English language; (c) investigated the treatment of a primary *Diagnostic and Statistical Manual of Mental Disorders (DSM)–III-*, *IV-*, or *V-*, or International Classification of Diseases-defined PD diagnosed through a validated diagnostic measure (e.g., Structured Clinical Interview for the Diagnosis of Axis-II Disorders [SCID-II]; First, Gibbon, Spitzer, Williams, & Benjamin, 1997; International Personality Disorder Examination; Loranger, 1995); (d) involved an adult (18+) sample; (e) included at least one PDT treatment group intended to treat PD; and took place in the context of a RCT that (f-i) compared PDT to another, nonpsychodynamic active treatment intended to produce a therapeutic effect over and above generic attention and support (e.g., cognitive behavioral therapy, DBT, psychopharmacology) or (f-ii) compared PDT to a control condition intended and expected to underperform against any uniquely therapeutic treatment (e.g., wait-list, a non-*bona fide* supportive counseling condition, treatment as usual).

[1] Two deviations from the protocol are noted. First, not enough studies were available in the prespecified follow-up blocks (short term and long term) to analyze; therefore, the longest available follow-up point from each study was grouped into a single posttreatment follow-up analysis. Second, reliability and consensus were hard to attain on moderator codes for relative strength of adherence checks between PDT and other active treatments, and these analyses were not conducted.

Outcomes and Effect Size Coding

A pool of three raters (John R. Keefe, Kevin S. McCarthy, and Ulrike Dinger) were randomly assigned to extract information to calculate effect sizes from identified studies. Every study effect size was extracted by two raters, and discrepancies were resolved by consultation of original publications or consensus between raters. When available, follow-up publications to the original trial were consulted for additional effect size information, which are noted in Table 1. Effect sizes were converted to Hedges' *g*. Effects were continuously coded such that negative *g* values represented a relative advantage for psychodynamic therapies versus other treatments (e.g., fewer core PD symptoms).

Our primary outcome measure was PD-specific symptomatology (e.g., SCID-II criteria counts, PD-specific measures, outcome measures not specifically designed for PD but conceptualized by the trial to conform to SCID-II diagnostic criteria, such as measures of verbal and direct aggression for borderline PD). For trials of BPD, suicidality was also considered a primary outcome measure. Secondary outcomes collected included Axis-I symptomatology (e.g., Symptom Checklist-90; Derogatis & Unger, 2010; Beck Depression Inventory; Beck, Steer, & Brown, 1996; Beck Anxiety Inventory; Beck, Epstein, Brown, & Steer, 1988), Inventory of Interpersonal Problems scores (Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988), psychosocial functioning (Global Assessment of Functioning; Aas, 2010; Social Adjustment Scale; Weissman & Bothwell, 1976), treatment dropout, and diagnostic remission from PD per a structured interview.

Quality Rating

The Randomized Controlled Trial Psychotherapy Quality Rating Scale (RCT-PQRS) was employed to assess the quality of the included trials (Kocsis et al., 2010). The scale is a psychometrically reliable and internally consistent measure generated through expert consultation with psychotherapy trial experts on the quality features that are relevant in assessing psychotherapy trials. The scale consists of 24 quality items assessing specific aspects of psychotherapy trial quality on a scale of 0–2, and an omnibus Item 25 (rated 0–7) offering a holistic sense for the quality of the study. For the total score, Kocsis and colleagues (2010) proposed a sum of 24 as signifying higher study quality, corresponding to an average of a "1" for each quality item. A copy of the measure is provided in the online supplemental materials.

A pool of three raters (Shelley F. McMain, Sigal Zilcha-Mano, and Zeynep Sahin) who did not participate in effect size coding were randomly assigned to quality rate studies, with two raters assigned per study. For the final ratings, interrater reliability (ICC[2,2]) for averaged quality scores was excellent for the sum of Items 1–24 (ICC = 0.91) and good for the omnibus Item 25 quality score (ICC = 0.87). The internal reliability of the item sum in this sample was also good ($\alpha = .84$). The item sum and omnibus quality rating were highly correlated with one another, $r(16) = 0.85$, $p < .001$.

Moderator Coding

The same raters who rated study quality for a given study also coded for moderators. Moderator coding was based on consensus between the two raters; discordant codes were returned to the given raters for them to work out. To avoid multiple testing, we examined moderators for effects only on core PD symptoms and general psychiatric symptoms.

Moderator variables specifically coded for were (a) for the PDT under study each of Wampold's original four criteria for establishing whether the therapy was *bona fide* (Wampold et al., 1997); (b) type of control treatment used (coded as either an inactive control like a wait-list, a control without guarantee of treatment like a treatment as usual [TAU] group; or an enhanced active control, typically a "manualized control" or a strong control with similar therapy hours as the enhanced treatment); (c) PD being studied; (d) number of therapy sessions total; and (e) frequency of therapy sessions per week.

Analyses

Random effects meta-analyses were conducted in the *R* statistical computing environment using the *R* package *metafor* (Viechtbauer, 2010). Our random effects analyses employed the Sidik-Jonkman estimator (Sidik & Jonkman, 2002) and the Knapp-Hartung adjustment to the standard errors (Hartung & Knapp, 2001), which in tandem produce more nominally accurate confidence intervals when including small studies and in the presence of between-study heterogeneity (IntHout, Ioannidis, & Borm, 2014).

Moderators of effect sizes were explored in a meta-regression framework. We explored the sensitivity of our findings by performing additional analyses in which we controlled for the magnitude of the pretreatment/baseline differences in symptom severity between the treatment groups (i.e., to what extent one group began treatment more impaired than another; see Table S4 in the online supplemental materials). To do so, we handicapped the standardized effect size of the baseline difference in severity from the termination effect size difference (e.g., if Treatment A was more severe at baseline by $g = 0.20$ relative to Treatment B, and at termination Treatment B had lower severity scores than A by $g = 0.20$, the adjusted effect would be $g = 0.00$). We also performed sensitivity tests for leaving any one study out, for comparisons that had at least four studies. For sensitivity analyses, we note any findings changing statistical significance at $p < .05$, effect size changes ($d \geq \pm 0.20$), or changes in heterogeneity (I^2 change $\geq 15\%$).

Heterogeneity of effect sizes was described using the I^2 statistic (Higgins, Thompson, Deeks, & Altman, 2003), with 25% reflecting low heterogeneity, 50% reflecting moderate heterogeneity, and 75%+ reflecting high heterogeneity. Heterogeneity indicates between-study differences in effect sizes that are not attributable to within-study sampling error.

To explore common patterns of effect sizes and sample size distributions that may indicate the presence of publication bias, both Egger's regression test (Egger, Davey Smith, Schneider, & Minder, 1997) and the Henmi and Copas method (Henmi & Copas, 2010) of detecting possible publication bias were employed. For comparisons of PDT with control groups, Rosenberg's variant of the fail-safe *N* was calculated (Rosenberg, 2005).

Power Analysis

We estimated the attained power of a random effects meta-analysis (Hedges & Pigott, 2001) based on our combined sample

Table 1
Characteristics of Included Studies

Study name	PDs studied	PDT treatment(s)	Active or control treatment(s)	Outcomes available	Follow-up time point (active Tx studies)	Total/omnibus quality score (1–7)
Abbass, Sheldon, Gyra, and Kalpin (2008)	Mixed PD	Avg. 27.7 sessions 1/wk ISTDP (n = 14)	Control: WLC (n = 13)	FUN, GEN, IIP	N/A	27/4
Amianto et al. (2011)	Borderline	40 sessions 1/wk SB-APP + eTAU (n = 17)	Control: eTAU (n = 16)	FUN, GEN, PD, SUC	N/A	23.5/4
Bateman and Fonagy (1999)	Borderline	Partial hospitalization 1.5 years MBT (n = 19)	Control: TAU (n = 19)	FUN, GEN, IIP, SUC	N/A	28/4.5
Bateman and Fonagy (2009)	Borderline	70 individual, 70 group 1/wk MBT (n = 71)	Control: eTAU (n = 63)	FUN, GEN, IIP, SUC	N/A	37/6.5
Clarkin, Levy, Lenzenweger, and Kernberg (2007)	Borderline	1 year 2/wk TFP; 1 year 1/wk SPP (n = 45)	Active: DBT (n = 17)	FUN, GEN, PD, SUC	None	32/6.5
Doering et al. (2010)	Borderline	1 year 2/wk TFP (n = 52)	Control: eTAU (n = 52)	FUN, GEN, PD, SUC	N/A	35.5/5.5
Emmelkamp et al. (2006)	Avoidant	20 sessions 1/wk Unmanualized PDT (n = 23)	Active: CBT (n = 18) Control: WLC (n = 16)	GEN, PD	6 months	17.5/2
Giesen-Bloo et al. (2006)	Borderline	3 years 2/wk TFP (n = 42)	Active: SFT (n = 44)	FUN, PD	None	34.5/5.5
Gregory et al. (2008)	Borderline (with alcohol use disorder)	1 year 1/wk DDP (n = 10)	Control: TAU (n = 9)	GEN, PD	N/A	32.5/4
Hellerstein et al. (1998)	Primarily Cluster C	40 sessions 1/wk STDP (n = 14)	Active: ST (n = 12)	GEN, IIP	6 months	26/3.5
Jørgensen et al. (2013, 2014)	Borderline	2 year 1 group 1 individual/wk MBT (n = 42)	Active: ST (n = 24)	FUN, GEN, IIP, PD	18 months	28.5/4.5
McMain et al. (2009, 2012)	Borderline	40 sessions GPM (n = 90)	Active: DBT (n = 90)	FUN, GEN, IIP, PD, SUC	24 months	41/6
Muran, Safran, Samstag, and Winston (2005)	Cluster C	30 sessions 1/wk BRT; STDP (n = 45)	Active: CBT (n = 29)	GEN, IIP, PD	6 months	29/3
Reneses et al. (2013)	Borderline	20 sessions 1/wk PRFP + TAU (n = 18)	Control: TAU (n = 26)	FUN, GEN, PD, SUC	N/A	21.5/3
Svartberg, Stiles, and Seltzer (2004)	Cluster C	40 sessions 1/wk AFT (n = 25)	Active: CBT (n = 25)	GEN, IIP, PD	24 months	33/5
Winston et al. (1994)	Primarily Cluster C	40 sessions 1/wk BAP; STDP (n = 55)	Control: WLC (n = 26)	FUN, GEN	N/A	20/3.5

Note. AFT = affect-focused psychodynamic therapy; Avg. = average; BAP = brief adaptive psychotherapy; BRT = brief relational therapy; CBT = cognitive behavioral therapy for personality disorders; DBT = dialectical behavior therapy; DDP = dynamic deconstructive therapy; eTAU = enhanced treatment as usual; FUN = functioning; GEN = general psychiatric symptoms; GPM = general psychodynamic psychiatric management; IIP = inventory of interpersonal problems; ISTDP = intensive short-term psychodynamic psychotherapy; MBT = mentalization-based treatment; PDs = personality disorder outcomes; PDT = psychodynamic therapies; PRFP = psychic representation focused psychotherapy; SB-APP = sequential brief Adlerian psychodynamic psychotherapy; SFT = schema-focused therapy; SPP = supportive psychodynamic therapy; ST = supportive therapy; STDP = short-term psychodynamic therapy; SUC = suicidality; TAU = treatment as usual; TFP = transference-focused psychotherapy; Tx = Treatment; WLC = wait-list control.

size for comparisons of PDT versus other active treatments and versus controls at treatment termination for core PD outcomes and general psychiatric symptoms, examining power to detect small ($g = 0.20$), medium ($g = 0.50$), and large ($g = 0.80$) effect sizes in cases of low (25%), medium (50%), and high (75%) between-study heterogeneity. Power was inadequate for small effects under heterogeneity (report in the online supplemental materials).

Results

Study Sample

Our final sample included 16 trials eligible to be meta-analyzed (see Figure 1 for PRISMA diagram; see Table 1 and Tables S1–S3 in the online supplemental materials for additional study characteristics). Eight studies compared PDT to another active treatment, and nine studies compared PDTs to a control condition (three wait-list controls, three TAUs, and three enhanced TAUs). A combined total of 585 patients were able to be included in active treatment comparisons at termination, and 519 patients were included in control group comparisons. Nine studies had follow-up assessments after treatment termination (range 6 months to 24 months; mean of 14 months). Ten trials treated BPD, four trials treated mixed exclusively or predominantly Cluster C PDs, one trial treated mixed PD including Cluster B PDs, and one trial treated avoidant PD (see Table 2 for full reporting of PD diagnosis rates from included trials). Only two studies (Bateman & Fonagy, 1999; Winston et al., 1994) were shared between this meta-analysis and that of Leichsenring and Leibing (2003).

The average trial was of adequate quality (total score $M = 29.2$, $SD = 6.5$; Item 25 score $M = 4.4$, $SD = 1.3$). Twelve out of 16 studies were above the 24-point normative threshold on the RCT-PQRS signifying a higher quality trial (Kocsis et al., 2010). At the level of a nonsignificant trend, trials treating primary BPD were of higher quality than other included trials (mean difference = +6), $t(15) = -1.94$, $p = .072$, $d = 1.01$. Total study quality had a small-to-medium size but nonsignificant correlation with study year, $r(16) = 0.27$, $p = .317$.

Control Group Comparisons

Core PD symptoms. In the unadjusted analysis, PDTs and control groups were not reliably distinguishable in treating core PD symptoms ($g = -0.25$; 95% confidence interval [CI; −0.82, 0.31], $SE = 0.20$, $p = .279$, $k = 5$), with a moderate level of heterogeneity ($I^2 = 56.9\%$). However, PDT groups were signifi-

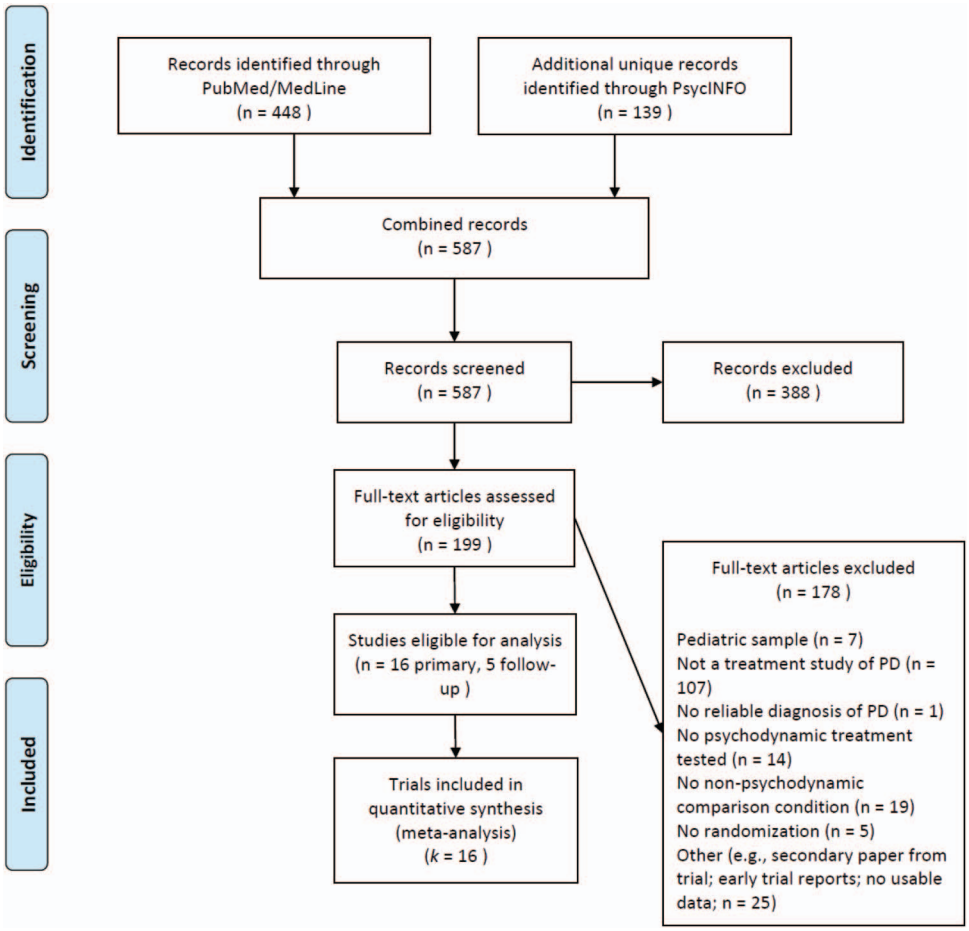

Figure 1. Flowchart of the selection of studies of psychodynamic therapy for the treatment of personality disorders (PD). See the online article for the color version of this figure.

Table 2
Personality Disorder Rates Reported in Each Study

Study name	Avoidant	Dependent	Obsessive–compulsive	Borderline	Narcissistic	Histrionic	Antisocial	Paranoid	PD-NOS
Abbass et al. (2008)	33.3%	7.4%	37.0%	44.4%	7.4%	3.7%	3.7%	18.5%	22.2%
Amianto et al. (2011)	NR	NR	NR	100%	NR	NR	NR	NR	NR
Bateman and Fonagy (1999)	NR	NR	NR	100%	NR	NR	NR	NR	NR
Bateman and Fonagy (2009)	20.9%	22.4%	3.0%	100%	18.7%	14.2%	27.6%	38.4%	NR
Clarkin et al. (2007)	NR	NR	NR	100%	NR	NR	NR	NR	NR
Doering et al. (2010)	NR	NR	NR	100%	NR	NR	0%	NR	NR
Emmelkamp et al. (2006)	100%	NR	NR	NR	NR	NR	NR	NR	NR
Giesen-Bloo et al. (2006)	NR	NR	NR	100%	NR	NR	0%	NR	NR
Gregory et al. (2008)	NR	NR	NR	100%	NR	NR	43.3%	NR	NR
Hellerstein et al. (1998)[c]	14.0%	4.0%	12.0%	0%	6%	6.0%	0%	10.0%	46.0%
Jørgensen et al. (2013, 2014)	23.4%	NR	NR	100%	NR	NR	0%	0%	NR
McMain et al. (2009, 2012)[a]	NR	NR	NR	100%	NR	NR	NR	NR	NR
Muran et al. (2005)[c]	22.0%	2.0%	10.0%	NR	NR	NR	NR	NR	66.0%
Reneses et al. (2013)	NR	NR	NR	100%	NR	NR	NR	NR	NR
Svartberg et al. (2004)	62.0%	34.0%	20.0%	0%	0%	0%	0%	0%	6%
Winston et al. (1994)[b]	NR	NR	NR	0%	0%	NR	NR	0%	29.6%

Note. PD-NOS = personality disorder not otherwise specified; NR = not reported. Rates do not add up to 100%, as patients could have multiple diagnoses. Rates of schizoid and schizotypal personality disorder are not reported, as they were negligibly present in studies assessing for them (<5%) or were exclusionary criteria.
[a] This study reported that 7.8% had a Cluster A disorder, 17.8% had an additional Cluster B disorder, and 40.6% had a Cluster C disorder. [b] This study reported that 18% had a Cluster B and 44% had a Cluster C disorder. [c] The study reported that these PD-NOS diagnoses primarily consisted of Cluster C traits.

cantly more impaired on core PD symptoms at baseline than control groups (weighted average $g = +0.35$, $p = .009$). After controlling for baseline differences on outcome measures there was a significant medium-sized effect in favor of PDT over control groups ($g = -0.63$; 95% CI [−0.87, −0.41], $SE = 0.08$, $p = .002$, fail-safe $n = 25$; see Figure 2). The adjustment according to baseline differences in outcome measures also resulted in substantively less heterogeneity ($I^2 = 11.4\%$). There was no clear single outlier in this adjusted group of studies. There was no evidence of publication bias.

Neither total quality ($B = -0.01$, $p = .490$) nor the omnibus Item 25 item ($B = -0.02$, $p = .808$) significantly predicted between-study effects. In addition, neither number ($B = -0.09$, $p = .779$) nor frequency of sessions ($B = -0.05$, $p = .812$) significantly predicted study effects. Number of *bona fide* criteria a PDT's study description met did not relate to study effects ($B = -0.10$, $p = .709$). It was not meaningful to examine whether effects differed between studies treating BPD and otherwise, as all but one study treated BPD, nor for strength of the control condition as examples of different ratings were sparse (one wait-list, three TAU, one enhanced TAU).

Suicidality. The unadjusted analysis revealed no significant differences between PDTs and control groups in improving suicidality ($g = -0.45$; 95% CI [−1.31, 0.40], $SE = 0.31$, $p = .217$, $k = 5$), with very high between-study heterogeneity ($I^2 = 85.0\%$). Again, patients in PDTs tended to begin treatment more suicidal than patients in control groups (weighted average $g = +0.33$, $p = .003$). Notably, in analyses controlling for baseline differences in severity of suicidality showed a significant, large effect size in favor of PDTs ($g = -0.79$; 95% CI [−1.38, −0.20], $SE = 0.21$, $p = .021$, fail-safe $n = 47$), with reduced but still high heterogeneity ($I^2 = 71.5\%$). Bateman and Fonagy (1999) was an outlier in favor of PDTs ($g = -1.89$). Removing the study from the analysis deflated the effect size to a medium-large effect ($g = -0.67$; 95% CI [−1.13, −0.20], $SE = 0.15$, $p = .020$) and reduced the heterogeneity to moderate ($I^2 = 40.1\%$).

There was no evidence of publication bias. Neither total study quality score ($B = -0.01$; 95% CI [−0.13, 0.11], $SE = 0.04$, $p = .773$) nor rated strength of the control condition related to outcomes (TAU vs. enhanced TAU), $F(1, 3) = 0.60$, $p = .494$.

General psychiatric symptoms. PDTs were reliably superior to control treatments in improving Axis-I/general psychiatric symptoms with a small-to-medium effect advantage ($g = -0.38$; 95% CI [−0.68, −0.08], $SE = 0.13$, $p = .019$, $k = 9$, fail-safe $n = 18$). However, there was a medium level of heterogeneity ($I^2 = 53.7\%$). Although there was no notable single-study outlier, this heterogeneity decreased if controlling for baseline severity of psychiatric symptoms ($g = -0.47$; 95% CI [−0.69, −0.25], $SE = 0.09$, $p = .001$, $I^2 = 34.87\%$).

Neither total quality ($B = -0.00$, $p = .950$) nor the omnibus Item 25 ($B = -0.03$, $p = .801$) score were predictive of treatment effects. In addition, neither session number ($B = -0.04$, $p = .852$) nor session frequency ($B = -0.04$, $p = .707$) predicted effects. Number of *bona fide* criteria did not predict effects ($B = -0.00$, $p = .998$). BPD trials did not significantly differ from other trials treating other disorders ($g = +0.09$, $p = .693$), but this finding is difficult to interpret, as BPD trials also all had active controls, whereas Cluster C trials all had inactive controls.

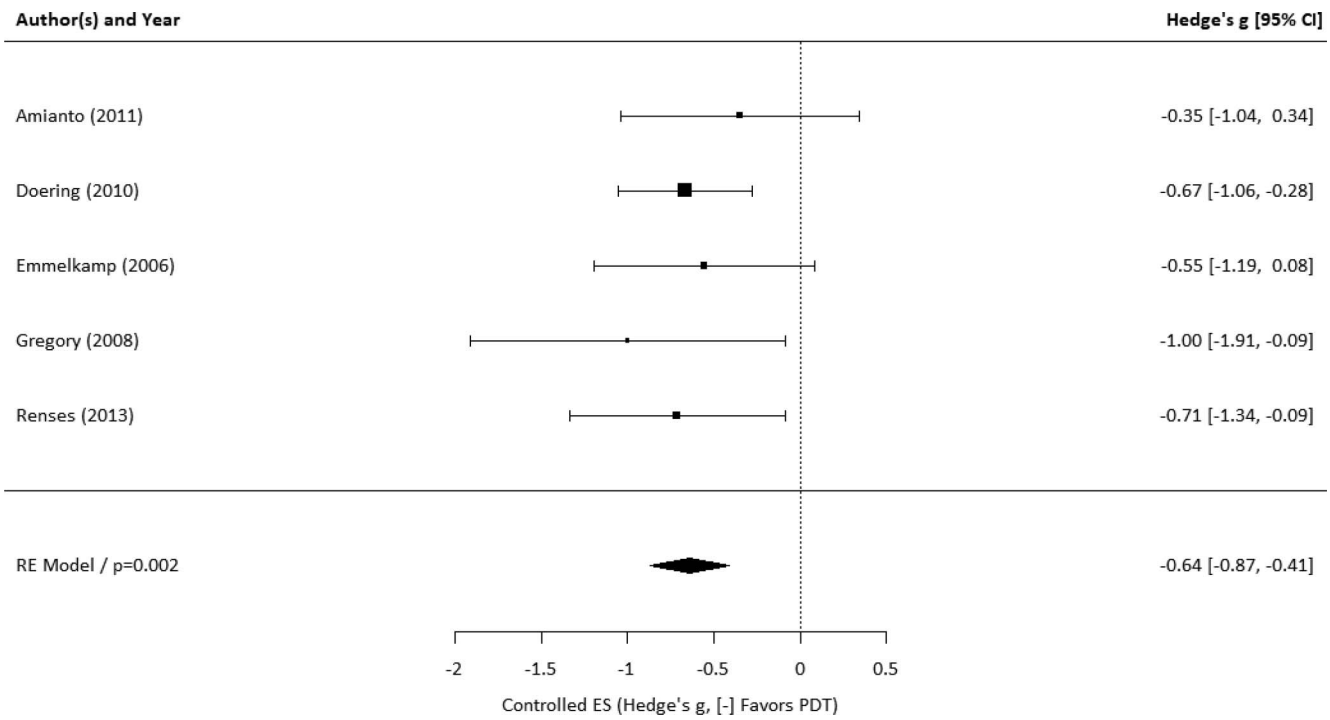

Figure 2. Effects of psychodynamic therapy compared with control groups on core personality disorder outcomes at treatment termination, adjusted for differences in baseline values. Psychodynamic treatments were significantly superior to control treatments with a medium-to-large effect size ($p = .002$, fail-safe $n = 25$). There were no clear single-study outliers, and there was no clear indication of publication bias.

Inventory of interpersonal problems. In improving inventory of interpersonal problems (IIP) scores, although each individual study showed a significant advantage of PDTs over control (g range = −0.69 to −2.22), the aggregate estimate was not statistically reliable in the meta-analysis due to the extremely high degree of heterogeneity introduced by a very large outlier among this small group of trials ($g = -1.25$; 95% CI [−3.22, 0.71], $SE = 0.46$, $p = .111$, $k = 3$, $I^2 = 86.2\%$).

Functioning. Patients significantly improved their psychosocial functioning in PDT as compared with control treatments ($g = -0.66$; 95% CI [−1.01, −0.32], $SE = 0.14$, $p = .003$, $k = 7$). This estimate exhibited a moderate degree of heterogeneity ($I^2 = 56.52\%$).

Active Treatment Comparisons: Termination

Core PD symptoms. PDTs and other psychotherapies evidenced comparable average outcomes in terms of treating core PD symptoms ($g = 0.05$; 95% CI [−0.25, 0.35], $SE = 0.12$, $p = .708$, $k = 7$; see Figure 3), with a moderate amount of between-study heterogeneity ($I^2 = 54.29\%$). The Giesen-Bloo et al. (2006) trial comparing transference-focused psychotherapy (TFP) to schema-focused therapy was a possible outlier, reducing heterogeneity somewhat when removed ($g = -0.04$; 95% CI [−0.31, 0.22], $SE = 0.10$, $p = .704$, $I^2 = 37.9\%$). There was no evidence for publication bias.

Quality as indicated by the total quality score ($B = -0.01$, $p = .730$) nor the omnibus Item 25 ($B = -0.04$, $p = .716$) did not significantly predict between-groups effect size. In addition, neither number ($B = 0.06$, $p = .861$) nor frequency of sessions ($B = 0.02$, $p = .933$) significantly predicted study effects. Number of *bona fide* criteria a study's PDT description met did not significantly predict effects ($B = -0.35$, $p = .163$). There were no differences in effect ($g = +0.02$, $p = .953$) between trials treating BPD ($k = 4$) and trials treating Cluster C PD ($k = 3$).

Remission from PD. A small number of studies reported on remission from PD diagnosis between PDTs and other treatments, and among these studies there were no significant differences (log odds of remission = 0.11; 95% CI [−2.77, 2.99], $SE = 0.67$, $p = .884$, $k = 3$). However, there was a high degree of between-study heterogeneity among this limited group of trials ($I^2 = 77.71\%$), which indicates these findings should be interpreted especially cautiously.

General psychiatric symptoms. In treating general psychiatric symptoms, PDTs and other active treatments had almost identical outcomes ($g = 0.00$; 95% CI [−0.22, 0.23], $SE = 0.09$, $p = .983$, $k = 7$), with a low-to-moderate degree of heterogeneity ($I^2 = 34.9\%$). The Hellerstein et al. (1998) trial was potentially an outlier, leading to a reduction in heterogeneity when removed ($g = -0.04$; 95% CI [−0.23, 0.15], $SE = 0.07$, $p = .636$, $I^2 = 20.02\%$).

Neither total quality ($B = -0.01$, $p = .427$) nor the omnibus Item 25 ($B = 0.08$, $p = .187$) score were predictive of treatment effects. PDTs with more sessions had a significantly stronger effects relative to those with fewer sessions ($B = -0.55$; 95% CI [−1.06, −0.04], $SE = 0.20$, $p = .040$, 52.50% heterogeneity accounted for by this moderator), and there was a trend toward the same effect for session frequency ($B = -0.35$; 95% CI [−0.77,

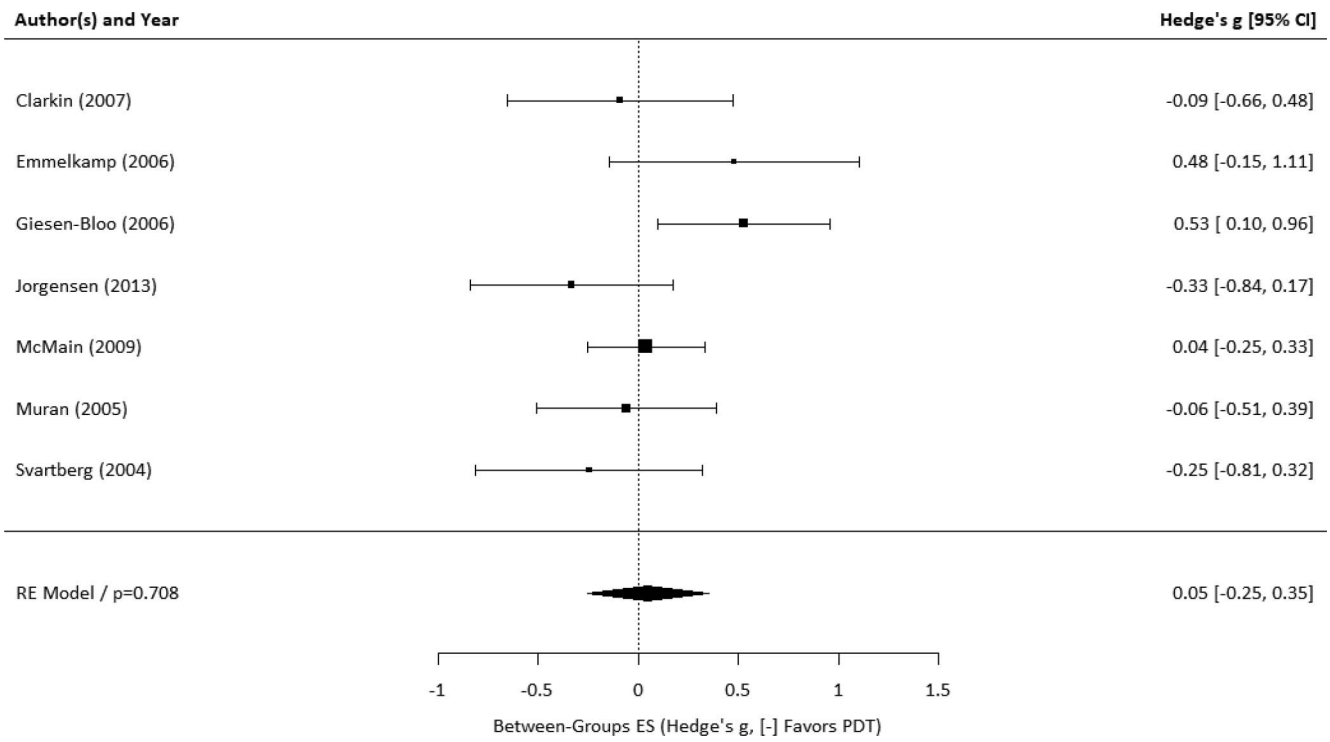

Figure 3. Effects of psychodynamic therapy compared with other *bona fide* active treatments for core PD outcomes at treatment termination. There was not a statistically reliable difference between PDT and other treatments, nor was there any indication of publication bias. The Giesen-Bloo et al. (2006) trial represented a possible outlier, but its removal did not change the statistical significance of the original finding.

0.08], $SE = 0.16$, $p = .073$). Number of *bona fide* criteria for the PDT therapy did not predict study effects ($B = -0.18$, $p = .384$). Studies treating BPD did not significantly differ from other trials ($d = -0.25$, $p = .199$).

Inventory of interpersonal problems. PDTs and other active treatments had very similar outcomes on the IIP ($g = -0.03$; 95% CI [−0.17, 0.10], $SE = 0.05$, $p = .541$, $k = 5$). There was no heterogeneity ($I^2 = 4.12\%$).

Functioning. PDTs and other active treatments had a comparable impact on improving psychosocial functioning ($g = 0.12$; 95% CI [−0.12, 0.36], $SE = 0.07$, $p = .202$, $k = 4$). There was minimal between-study heterogeneity ($I^2 = 2.34\%$).

Dropout. PDTs and other active treatments had comparable dropout (log RR = 0.01; 95% CI [−0.41, 0.43], $SE = 0.18$, $p = .970$, $k = 8$), but with medium-to-high heterogeneity in relative dropout between studies ($I^2 = 67.7\%$). Emmelkamp et al. (2006) was an outlier in favor of PDT relative to CT for dropout, and when removed heterogeneity decreased notably (log RR = 0.04; 95% CI [−0.32, 0.40], $SE = 0.15$, $p = .785$, $I^2 = 45.5\%$). Trials treating BPD were not distinguished from other trials on relative dropout, $F(1, 6) = 0.08$, $p = .790$.

Active Treatment Comparisons: Follow-Up

Core PD symptoms. There was not differential improvement in core PD symptoms at follow-up between PDT and other active treatments ($g = 0.00$; 95% CI [−0.48, 0.49], $SE = 0.17$, $p = .996$, $k = 5$), albeit with a medium-high level of heterogeneity ($I^2 = 63.66\%$). This heterogeneity was almost eliminated by removing the Emmelkamp et al. (2006) study, which resulted in a nonsignificant trend of a small advantage for PDTs at follow-up ($g = -0.18$; 95% CI [−0.38, 0.03], $SE = 0.06$, $p = .072$, $I^2 = 5.08\%$).

General psychiatric symptoms. General psychiatric symptom outcomes were not significantly different at follow-up between PDTs and other active treatments ($g = -0.14$; 95% CI [−0.43, 0.16], $SE = 0.11$, $p = .267$, $k = 5$), with a low-to-moderate degree of heterogeneity ($I^2 = 32.19$). The Hellerstein et al. (1998) trial was potentially an outlier, leading to a reduction in heterogeneity when removed and a nonsignificant trend toward a small advantage for PDT ($g = -0.20$; 95% CI [−0.44, 0.05], $SE = 0.08$, $p = .085$, $I^2 = 12.21$).

Inventory of interpersonal problems. Overall, PDTs did not significantly differ from other active treatments at follow-up in terms of IIP scores ($g = -0.17$; 95% CI [−0.42, 0.08], $SE = 0.09$, $p = .136$, $k = 5$), with a low amount of heterogeneity ($I^2 = 27.33\%$). Removing Hellerstein et al. (1998) resulted in a lack of heterogeneity and a significant, small effect size estimate in favor of PDTs over comparison treatments ($g = -0.23$; 95% CI [−0.28, −0.17], $SE = 0.02$, $p = .001$, $I^2 = 0.05\%$).

Discussion

For all outcomes, PDTs did not significantly differ in efficacy from other *bona fide* treatments in treating PDs (predominantly BPD and mixed Cluster C disorders). Nonsignificant differences

were all small in effect size. This parity continued into posttreatment follow-up. In addition, especially after adjustment for baseline symptom differences PDTs outperformed control conditions with medium effect sizes, with the exception of interpersonal problems wherein a reliable estimate could not be made due to very high between-study heterogeneity. In most cases the advantage of PDT over controls represented comparisons versus an active control ($k = 6$) rather than a wait-list control ($k = 3$). Furthermore, trials treating BPD versus other PDs did not evidence significantly different effect sizes, although this comparison was limited in power by the lower number of studies not targeting BPD ($k = 6$).

However, trials targeting Cluster C PD all used inactive control groups or head-to-head comparisons with other treatments that have also not been compared with active controls (e.g., cognitive behavioral therapy), versus trials treating BPD that all used active control groups. Moreover, there was a trend toward these trials having lower quality scores than trials for BPD ($d = -1.01$). Although PDTs consistently outperformed wait-list controls in treating Cluster C PD samples, it is unclear whether PDTs (or indeed the other active treatments tested in these trials) have unique therapeutic benefit over less specific approaches for these conditions.

Importantly, PDTs evidenced medium-to-large effect size advantages over active control treatments in improving suicidality among patients with BPD. This finding is commensurate with a meta-analysis of PDTs finding significant benefits versus control treatments on suicide attempts and self-harm incidents across disorders (Briggs et al., 2019). In the two trials in which a PDT for BPD was compared with DBT (Clarkin, Levy, Lenzenweger, & Kernberg, 2007; McMain, Guimond, Streiner, Cardish, & Links, 2012; McMain et al., 2009), PDTs did not significantly differ from DBT in improving suicidality, but more studies are needed to run a formal meta-analysis. PDT for PD exhibited similar rates of dropout as compared with other active conditions for PD, suggesting that PDT was as acceptable to patients as other treatments for PD.

Average study quality was acceptable (with 75% of studies meeting the normative threshold), and in no comparison did effects differ reliably based on study quality. This follows null findings using the RCT-PQRS to examine the relationship of study quality to outcomes in PDT trials of anxiety disorders (Keefe, McCarthy, Dinger, Zilcha-Mano, & Barber, 2014) and no clear pattern of effects on the basis of quality from a qualitative review of PDT outcomes across different disorders (Gerber et al., 2011). There was also no clear evidence for publication bias.

Only one moderator was significant—in the treatment of general psychiatric symptoms (but not core PD symptoms), PDTs exhibited stronger effects relative to other active treatments when it was conducted over more sessions. This finding may track previously reported advantages of long-term PDT over shorter treatments for complex mental disorders (Leichsenring & Rabung, 2011); however, it should be interpreted cautiously given the lack of replication against control groups or in core personality disorder symptoms as an outcome. Number of *bona fide* criteria (Wampold et al., 1997) the PDT therapy met per the study description also did not reliably predict effects.

Limitations

Although the meta-analysis attained adequate power to detect medium to large effect differences between PDTs and other treatments at treatment termination, our study was not powered to detect small effects under conditions of any effect heterogeneity. It is possible that small effect size differences exist between PDTs and other therapies that were not detectable. Studies comparing PDT to control groups generally did not report posttreatment follow-up, although those that did (all for BPD) generally found maintenance of PDT superiority at follow-up (Bateman & Fonagy, 2001, 2008; Gregory, DeLucia-Deranja, & Mogle, 2010), although with overall a small sample. Further studies comparing PDT to controls over follow-up would be warranted, particularly against more rigorous control groups (Bateman & Fonagy, 2009). More well-powered studies of PDT are also needed (Bateman & Fonagy, 2009; McMain et al., 2009).

There were also no studies addressing Cluster A PDs, only one study examining effects for a specific Cluster C PD (Emmelkamp et al., 2006), no studies for histrionic or antisocial PD, and no studies addressing narcissistic PD in either its *DSM–V* definition or as a personality spectrum encompassing narcissistic entitlement, grandiosity, and vulnerability (Krizan & Herlache, 2018). Among trials aiming at treating Cluster C PD ($k = 6$), avoidant PD and personality disorder not otherwise specified with Cluster C traits were the most commonly diagnosed (see Table 2), with underrepresentation of both obsessive–compulsive and dependent PDs. The specific effects of PDT on individual Cluster C disorders remain unclear, a general limitation of the Cluster C PD treatment literature (e.g., Bamelis, Evers, Spinhoven, & Arntz, 2014). Moreover, despite a need for treatments that may engage more narcissistic patients who are liable to drop out of treatments as usual (Ellison, Levy, Cain, Ansell, & Pincus, 2013; Hilsenroth, Holdwick, Castlebury, & Blais, 1998), no targeted PDT (or any therapy) for narcissism has ever been tested (Caligor, Levy, & Yeomans, 2015), a significant clinical gap.

Although several of our outcome categories were relatively homogenous in terms of measurement instrument, trials used different outcome measures to assess core PD symptoms. For example, among the eight BPD trials included in this meta-analysis reporting on core PD outcomes, six different assessment strategies for BPD symptoms were employed. The range of different measurement instruments employed could have contributed to observed between-study effect heterogeneity and render cross-comparisons of effect sizes potentially less interpretable.

In addition, five of the 16 studies (four vs. control) did not include a measure intended to assess core PD symptoms, limiting their contribution to the primary question of our meta-analysis. The smaller number of control group studies including core PD outcomes may have contributed to the imbalance of baseline severity in core PD symptoms detected (PDT > controls in severity). With fewer studies ($k = 5$ for controls vs. 7 for active treatments), there is a lower probability that the average baseline severity will be balanced by randomization, especially when individual studies have smaller sample sizes. Overall, PDT did evidence more improvements on core PD symptoms and suicidality than control groups, although this difference was obscured when the starting values of the two conditions were not adjusted in the analysis.

Future Directions

Study-level meta-analyses cannot inform as to what specific treatments for PDs may be most appropriate for particular PD patients. We strongly endorse the point of view that no single treatment for PD will work adequately for all patients. Specifically, given the generally similar effects observed between PDTs and other therapies in BPD, moving to questions of "what works for whom" in its treatment may be warranted. Object-relations psychotherapy (similar to TFP) has been found to be particularly effective compared with DBT and TAU among BPD patients with relatively higher baseline psychosocial functioning (Sahin et al., 2018). Conversely, one randomized trial suggests that mentalization-based treatment (MBT) may be more effective than usual care among *more* impaired patients (Bateman & Fonagy, 2013). More agreeable patients may form especially good therapeutic alliances facilitating treatment effects in DBT as compared with psychodynamic general psychiatric management (Hirsh, Quilty, Bagby, & McMain, 2012). As multiple effective treatments for PDs are identified, sophisticated moderator research may help inform personalized treatment decisions.

Finally, increasingly PD is understood psychometrically as a continuous spectrum indicating increasing deficits in self- and other-functioning common across PD diagnoses, as emphasized in the *DSM–V* alternative model for PDs and the International Classification of Diseases 11th Revision system for diagnosing PD (Morey, Benson, Busch, & Skodol, 2015; Strickland et al., 2019). PDTs focus on different aspects of self- and other-functioning that are also conceptualized on a severity continuum, such as deficits in epistemic trust and mentalization in MBT (Fonagy, Luyten, & Allison, 2015) and personality organization in TFP (Stern et al., 2010). Studies examining the differential effects of PDTs on common features of PD versus specific personality constellations (e.g., obsessionality; Sharp et al., 2015; Wright, Hopwood, Skodol, & Morey, 2016) may more precisely identify the impact of PDTs on specific PDs and may point toward potential transdiagnostic effects of these treatments.

Conclusion

PDTs are potentially effective in the treatment of borderline PD, where PDTs consistently outperformed active control conditions. The evidence base for PDT for Cluster C PDs is less clear, as all tested control groups were wait-lists. There is not strong indication PDTs are any better or worse on average than other targeted treatments for borderline and mixed Cluster C PDs.

References

References beginning with an asterisk indicate the study is included in the meta-analysis.

Aas, I. H. M. (2010). Global Assessment of Functioning (GAF): Properties and frontier of current knowledge. *Annals of General Psychiatry, 9,* 20. http://dx.doi.org/10.1186/1744-859X-9-20

*Abbass, A., Sheldon, A., Gyra, J., & Kalpin, A. (2008). Intensive short-term dynamic psychotherapy for *DSM–IV* personality disorders: A randomized controlled trial. *Journal of Nervous and Mental Disease, 196,* 211–216. http://dx.doi.org/10.1097/NMD.0b013e3181662ff0

*Amianto, F., Ferrero, A., Pierò, A., Cairo, E., Rocca, G., Simonelli, B., . . . Fassino, S. (2011). Supervised team management, with or without structured psychotherapy, in heavy users of a mental health service with borderline personality disorder: A two-year follow-up preliminary randomized study. *BMC Psychiatry, 11,* 181. http://dx.doi.org/10.1186/1471-244X-11-181

Ansell, E. B., Pinto, A., Edelen, M. O., Markowitz, J. C., Sanislow, C. A., Yen, S., . . . Grilo, C. M. (2011). The association of personality disorders with the prospective 7-year course of anxiety disorders. *Psychological Medicine, 41,* 1019–1028. http://dx.doi.org/10.1017/S0033291710001777

Ansell, E. B., Sanislow, C. A., McGlashan, T. H., & Grilo, C. M. (2007). Psychosocial impairment and treatment utilization by patients with borderline personality disorder, other personality disorders, mood and anxiety disorders, and a healthy comparison group. *Comprehensive Psychiatry, 48,* 329–336. http://dx.doi.org/10.1016/j.comppsych.2007.02.001

Auchincloss, E. L. (2015). *The psychoanalytic model of mind*. Arlington, VA: American Psychiatric Publishing.

Bamelis, L. L., Evers, S. M., Spinhoven, P., & Arntz, A. (2014). Results of a multicenter randomized controlled trial of the clinical effectiveness of schema therapy for personality disorders. *The American Journal of Psychiatry, 171,* 305–322. http://dx.doi.org/10.1176/appi.ajp.2013.12040518

Barber, J. P., Muran, J. C., McCarthy, K. S., & Keefe, J. R. (2013). Research on dynamic therapies. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed., pp. 443–494). New York, NY: Wiley.

*Bateman, A., & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder: A randomized controlled trial. *The American Journal of Psychiatry, 156,* 1563–1569. http://dx.doi.org/10.1176/ajp.156.10.1563

Bateman, A., & Fonagy, P. (2001). Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: An 18-month follow-up. *The American Journal of Psychiatry, 158,* 36–42. http://dx.doi.org/10.1176/appi.ajp.158.1.36

Bateman, A., & Fonagy, P. (2008). 8-year follow-up of patients treated for borderline personality disorder: Mentalization-based treatment versus treatment as usual. *The American Journal of Psychiatry, 165,* 631–638. http://dx.doi.org/10.1176/appi.ajp.2007.07040636

*Bateman, A., & Fonagy, P. (2009). Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *The American Journal of Psychiatry, 166,* 1355–1364. http://dx.doi.org/10.1176/appi.ajp.2009.09040539

Bateman, A., & Fonagy, P. (2013). Impact of clinical severity on outcomes of mentalisation-based treatment for borderline personality disorder. *The British Journal of Psychiatry, 203,* 221–227. http://dx.doi.org/10.1192/bjp.bp.112.121129

Bateman, A., & Fonagy, P. (2016). *Mentalization based treatment for personality disorders: A practical guide*. Oxford, United Kingdom: Oxford University Press. http://dx.doi.org/10.1093/med:psych/9780199680375.001.0001

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology, 56,* 893–897. http://dx.doi.org/10.1037/0022-006X.56.6.893

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory-II (BDI-II)*. San Antonio, TX: Psychological Corporation.

Briggs, S., Netuveli, G., Gould, N., Gkaravella, A., Gluckman, N. S., Kangogyere, P., . . . Lindner, R. (2019). The effectiveness of psychoanalytic/psychodynamic psychotherapy for reducing suicide attempts and self-harm: Systematic review and meta-analysis. *The British Journal of Psychiatry, 214,* 320–328. http://dx.doi.org/10.1192/bjp.2019.33

Caligor, E., Levy, K. N., & Yeomans, F. E. (2015). Narcissistic personality disorder: Diagnostic and clinical challenges. *The American Journal of Psychiatry, 172,* 415–422. http://dx.doi.org/10.1176/appi.ajp.2014.14060723

*Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2007). Evaluating three treatments for borderline personality disorder: A multiwave study. *The American Journal of Psychiatry, 164,* 922–928. http://dx.doi.org/10.1176/ajp.2007.164.6.922

Cristea, I. A., Gentili, C., Cotet, C. D., Palomba, D., Barbui, C., & Cuijpers, P. (2017). Efficacy of psychotherapies for borderline personality disorder: A systematic review and meta-analysis. *Journal of the American Medical Association Psychiatry, 74,* 319–328. http://dx.doi.org/10.1001/jamapsychiatry.2016.4287

Cuijpers, P., Weitz, E., Cristea, I. A., & Twisk, J. (2017). Pre-post effect sizes should be avoided in meta-analyses. *Epidemiology and Psychiatric Sciences, 26,* 364–368. http://dx.doi.org/10.1017/S2045796016000809

Derogatis, L. R., & Unger, R. (2010). Symptom Checklist-90–Revised. In I. B. Weiner & W. E. Craighead (Eds.), *The Corsini encyclopedia of psychology* (pp. 81–84). Hoboken, NJ: Wiley. http://dx.doi.org/10.1002/9780470479216.corpsy0970

*Doering, S., Hörz, S., Rentrop, M., Fischer-Kern, M., Schuster, P., Benecke, C., . . . Buchheim, P. (2010). Transference-focused psychotherapy v. treatment by community psychotherapists for borderline personality disorder: Randomised controlled trial. *The British Journal of Psychiatry, 196,* 389–395. http://dx.doi.org/10.1192/bjp.bp.109.070177

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315,* 629–634. http://dx.doi.org/10.1136/bmj.315.7109.629

Ellison, W. D., Levy, K. N., Cain, N. M., Ansell, E. B., & Pincus, A. L. (2013). The impact of pathological narcissism on psychotherapy utilization, initial symptom severity, and early-treatment symptom change: A naturalistic investigation. *Journal of Personality Assessment, 95,* 291–300. http://dx.doi.org/10.1080/00223891.2012.742904

*Emmelkamp, P. M., Benner, A., Kuipers, A., Feiertag, G. A., Koster, H. C., & van Apeldoorn, F. J. (2006). Comparison of brief dynamic and cognitive-behavioural therapies in avoidant personality disorder. *The British Journal of Psychiatry, 189,* 60–64. http://dx.doi.org/10.1192/bjp.bp.105.012153

First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *Structured Clinical Interview for DSM–IV Axis II personality disorders (SCID-II)*. Washington, DC: American Psychiatric Press.

Fonagy, P., Luyten, P., & Allison, E. (2015). Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment. *Journal of Personality Disorders, 29,* 575–609. http://dx.doi.org/10.1521/pedi.2015.29.5.575

Gerber, A. J., Kocsis, J. H., Milrod, B. L., Roose, S. P., Barber, J. P., Thase, M. E., . . . Leon, A. C. (2011). A quality-based review of randomized controlled trials of psychodynamic psychotherapy. *The American Journal of Psychiatry, 168,* 19–28. http://dx.doi.org/10.1176/appi.ajp.2010.08060843

*Giesen-Bloo, J., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T., . . . Arntz, A. (2006). Outpatient psychotherapy for borderline personality disorder: Randomized trial of schema-focused therapy vs transference-focused psychotherapy. *Archives of General Psychiatry, 63,* 649–658. http://dx.doi.org/10.1001/archpsyc.63.6.649

*Gregory, R. J., Chlebowski, S., Kang, D., Remen, A. L., Soderberg, M. G., Stepkovitch, J., & Virk, S. (2008). A controlled trial of psychodynamic psychotherapy for co-occurring borderline personality disorder and alcohol use disorder. *Psychotherapy: Theory, Research, and Practice, 45,* 28–41. http://dx.doi.org/10.1037/0033-3204.45.1.28

Gregory, R. J., DeLucia-Deranja, E., & Mogle, J. A. (2010). Dynamic deconstructive psychotherapy versus optimized community care for borderline personality disorder co-occurring with alcohol use disorders: A 30-month follow-up. *Journal of Nervous and Mental Disease, 198,* 292–298. http://dx.doi.org/10.1097/NMD.0b013e3181d6172d

Grilo, C. M., Stout, R. L., Markowitz, J. C., Sanislow, C. A., Ansell, E. B., Skodol, A. E., . . . McGlashan, T. H. (2010). Personality disorders predict relapse after remission from an episode of major depressive disorder: A 6-year prospective study. *The Journal of Clinical Psychiatry, 71,* 1629–1635. http://dx.doi.org/10.4088/JCP.08m04200gre

Gunderson, J. G., & Links, P. F. (2014). *Handbook of good psychiatric management for borderline personality disorder*. Arlington, VA: American Psychiatric Publishing.

Hartung, J., & Knapp, G. (2001). On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Statistics in Medicine, 20,* 1771–1782. http://dx.doi.org/10.1002/sim.791

Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods, 6,* 203–217. http://dx.doi.org/10.1037/1082-989X.6.3.203

*Hellerstein, D. J., Rosenthal, R. N., Pinsker, H., Samstag, L. W., Muran, J. C., & Winston, A. (1998). A randomized prospective study comparing supportive and dynamic therapies. Outcome and alliance. *The Journal of Psychotherapy Practice and Research, 7,* 261–271.

Henmi, M., & Copas, J. B. (2010). Confidence intervals for random effects meta-analysis and robustness to publication bias. *Statistics in Medicine, 29,* 2969–2983. http://dx.doi.org/10.1002/sim.4029

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557–560. http://dx.doi.org/10.1136/bmj.327.7414.557

Hilsenroth, M. J., Holdwick, D. J., Jr., Castlebury, F. D., & Blais, M. A. (1998). The effects of *DSM–IV* cluster B personality disorder symptoms on the termination and continuation of psychotherapy. *Psychotherapy: Theory, Research, and Practice, 35,* 163–176. http://dx.doi.org/10.1037/h0087845

Hirsh, J. B., Quilty, L. C., Bagby, R. M., & McMain, S. F. (2012). The relationship between agreeableness and the development of the working alliance in patients with borderline personality disorder. *Journal of Personality Disorders, 26,* 616–627. http://dx.doi.org/10.1521/pedi.2012.26.4.616

Høglend, P., & Hagtvet, K. (2019). Change mechanisms in psychotherapy: Both improved insight and improved affective awareness are necessary. *Journal of Consulting and Clinical Psychology, 87,* 332–344. http://dx.doi.org/10.1037/ccp0000381

Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56,* 885–892. http://dx.doi.org/10.1037/0022-006X.56.6.885

Huang, Y., Kotov, R., de Girolamo, G., Preti, A., Angermeyer, M., Benjet, C., . . . Kessler, R. C. (2009). *DSM–IV* personality disorders in the WHO World Mental Health Surveys. *The British Journal of Psychiatry, 195,* 46–53. http://dx.doi.org/10.1192/bjp.bp.108.058552

IntHout, J., Ioannidis, J. P., & Borm, G. F. (2014). The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Medical Research Methodology, 14,* 25. http://dx.doi.org/10.1186/1471-2288-14-25

Johansen, P. O., Krebs, T. S., Svartberg, M., Stiles, T. C., & Holen, A. (2011). Change in defense mechanisms during short-term dynamic and cognitive therapy in patients with cluster C personality disorders. *Journal of Nervous and Mental Disease, 199,* 712–715. http://dx.doi.org/10.1097/NMD.0b013e318229d6a7

Johansson, P., Høglend, P., Ulberg, R., Amlo, S., Marble, A., Bøgwald, K. P., . . . Heyerdahl, O. (2010). The mediating role of insight for long-term improvements in psychodynamic therapy. *Journal of Consulting and Clinical Psychology, 78,* 438–448. http://dx.doi.org/10.1037/a0019245

*Jørgensen, C. R., Bøye, R., Andersen, D., Døssing Blaabjerg, A. H., Freund, C., Jordet, H., & Kjølbye, M. (2014). Eighteen months post-treatment naturalistic follow-up study of mentalization-based therapy and supportive group treatment of borderline personality disorder: Clin-

ical outcomes and functioning. *Nordic Psychology, 66,* 254–273. http://dx.doi.org/10.1080/19012276.2014.963649

*Jørgensen, C. R., Freund, C., Bøye, R., Jordet, H., Andersen, D., & Kjølbye, M. (2013). Outcome of mentalization-based and supportive psychotherapy in patients with borderline personality disorder: A randomized trial. *Acta Psychiatrica Scandinavica, 127,* 305–317. http://dx.doi.org/10.1111/j.1600-0447.2012.01923.x

Kallestad, H., Valen, J., McCullough, L., Svartberg, M., Høglend, P., & Stiles, T. C. (2010). The relationship between insight gained during therapy and long-term outcome in short-term dynamic psychotherapy and cognitive therapy for cluster C personality disorders. *Psychotherapy Research, 20,* 526–534. http://dx.doi.org/10.1080/10503307.2010.492807

Keefe, J. R. (2015). Heightened risk of false positives in a network meta-analysis of social anxiety. *The Lancet Psychiatry, 2,* 292–293. http://dx.doi.org/10.1016/S2215-0366(15)00043-7

Keefe, J. R., & DeRubeis, R. J. (2019). Changing character: A narrative review of personality change in psychotherapies for personality disorder. *Psychotherapy Research, 29,* 752–769. http://dx.doi.org/10.1080/10503307.2018.1425930

Keefe, J. R., McCarthy, K. S., Dinger, U., Zilcha-Mano, S., & Barber, J. P. (2014). A meta-analytic review of psychodynamic therapies for anxiety disorders. *Clinical Psychology Review, 34,* 309–323. http://dx.doi.org/10.1016/j.cpr.2014.03.004

Keefe, J. R., Milrod, B. L., Gallop, R., Barber, J. P., & Chambless, D. L. (2018). What is the effect on comorbid personality disorder of brief panic-focused psychotherapy in patients with panic disorder? *Depression and Anxiety, 35,* 239–247. http://dx.doi.org/10.1002/da.22708

Keefe, J. R., Webb, C. A., & DeRubeis, R. J. (2016). In cognitive therapy for depression, early focus on maladaptive beliefs may be especially efficacious for patients with personality disorders. *Journal of Consulting and Clinical Psychology, 84,* 353–364. http://dx.doi.org/10.1037/ccp0000071

Kocsis, J. H., Gerber, A. J., Milrod, B., Roose, S. P., Barber, J., Thase, M. E., . . . Leon, A. C. (2010). A new scale for assessing the quality of randomized clinical trials of psychotherapy. *Comprehensive Psychiatry, 51,* 319–324. http://dx.doi.org/10.1016/j.comppsych.2009.07.001

Krizan, Z., & Herlache, A. D. (2018). The Narcissism Spectrum Model: A synthetic view of narcissistic personality. *Personality and Social Psychology Review, 22,* 3–31. http://dx.doi.org/10.1177/1088868316685018

Leichsenring, F., & Leibing, E. (2003). The effectiveness of psychodynamic therapy and cognitive behavior therapy in the treatment of personality disorders: A meta-analysis. *The American Journal of Psychiatry, 160,* 1223–1232. http://dx.doi.org/10.1176/appi.ajp.160.7.1223

Leichsenring, F., & Rabung, S. (2011). Long-term psychodynamic psychotherapy in complex mental disorders: Update of a meta-analysis. *The British Journal of Psychiatry, 199,* 15–22. http://dx.doi.org/10.1192/bjp.bp.110.082776

Lenzenweger, M. F., Lane, M. C., Loranger, A. W., & Kessler, R. C. (2007). *DSM–IV* personality disorders in the National Comorbidity Survey Replication. *Biological Psychiatry, 62,* 553–564. http://dx.doi.org/10.1016/j.biopsych.2006.09.019

Lilliengren, P. (2018, December 27). *Comprehensive compilation of randomized controlled trials (RCTs) involving psychodynamic treatments and interventions*. Retrieved from https://www.researchgate.net/publication/317335876_Comprehensive_compilation_of_randomized_controlled_trials_RCTs_involving_psychodynamic_treatments_and_interventions

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York, NY: Guilford Press.

Loranger, A. W. (1995). *International Personality Disorders Examination (IPDE) Manual*. White Plains, NY: Weill-Cornell Medical Centre.

Luborsky, L. (1984). *Principles of psychoanalytic psychotherapy: Manual for supportive-expressive treatment*. New York, NY: Basic Books.

Magnavita, J. J., Levy, K. N., Critchfield, K. L., & Lebow, J. L. (2010). Ethical considerations in treatment of personality dysfunction: Using evidence, principles, and clinical judgment. *Professional Psychology, Research and Practice, 41,* 64–74. http://dx.doi.org/10.1037/a0017733

Markowitz, J. C., Skodol, A. E., Petkova, E., Cheng, J., Sanislow, C. A., Grilo, C. M., . . . McGlashan, T. H. (2007). Longitudinal effects of personality disorders on psychosocial functioning of patients with major depressive disorder. *The Journal of Clinical Psychiatry, 68,* 186–193. http://dx.doi.org/10.4088/JCP.v68n0202

McCullough, L., Kuhn, N., Andrews, S., Kaplan, A., Wolf, J., & Hurley, C. L. (2003). *Treating affect phobia: A manual for short-term dynamic psychotherapy*. New York, NY: Guilford Press.

*McMain, S. F., Guimond, T., Streiner, D. L., Cardish, R. J., & Links, P. S. (2012). Dialectical behavior therapy compared with general psychiatric management for borderline personality disorder: Clinical outcomes and functioning over a 2-year follow-up. *The American Journal of Psychiatry, 169,* 650–661. http://dx.doi.org/10.1176/appi.ajp.2012.11091416

*McMain, S. F., Links, P. S., Gnam, W. H., Guimond, T., Cardish, R. J., Korman, L., & Streiner, D. L. (2009). A randomized trial of dialectical behavior therapy versus general psychiatric management for borderline personality disorder. *The American Journal of Psychiatry, 166,* 1365–1374. http://dx.doi.org/10.1176/appi.ajp.2009.09010039

Morey, L. C., Benson, K. T., Busch, A. J., & Skodol, A. E. (2015). Personality disorders in *DSM–5*: Emerging research on the alternative model. *Current Psychiatry Reports, 17,* 24. http://dx.doi.org/10.1007/s11920-015-0558-0

*Muran, J. C., Safran, J. D., Samstag, L. W., & Winston, A. (2005). Evaluating an alliance-focused treatment for personality disorders. *Psychotherapy: Theory, Research, and Practice, 42,* 532–545. http://dx.doi.org/10.1037/0033-3204.42.4.532

Newton-Howes, G., Tyrer, P., Johnson, T., Mulder, R., Kool, S., Dekker, J., & Schoevers, R. (2014). Influence of personality on the outcome of treatment in depression: Systematic review and meta-analysis. *Journal of Personality Disorders, 28,* 577–593. http://dx.doi.org/10.1521/pedi_2013_27_070

Perry, J. C., & Bond, M. (2012). Change in defense mechanisms during long-term dynamic psychotherapy and five-year outcome. *The American Journal of Psychiatry, 169,* 916–925. http://dx.doi.org/10.1176/appi.ajp.2012.11091403

Perry, J. C., & Bond, M. (2017). Addressing defenses in psychotherapy to improve adaptation. *Psychoanalytic Inquiry, 37,* 153–166. http://dx.doi.org/10.1080/07351690.2017.1285185

*Reneses, B., Galián, M., Serrano, R., Figuera, D., Fernandez Del Moral, A., López-Ibor, J. J., . . . Trujillo, M. (2013). A new time limited psychotherapy for BPD: Preliminary results of a randomized and controlled trial. *Actas Españolas de Psiquiatría, 41,* 139–148.

Rosenberg, M. S. (2005). The file-drawer problem revisited: A general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution; International Journal of Organic Evolution, 59,* 464–468. http://dx.doi.org/10.1111/j.0014-3820.2005.tb01004.x

Sahin, Z., Vinnars, B., Gorman, B. S., Wilczek, A., Åsberg, M., & Barber, J. P. (2018). Clinical severity as a moderator of outcome in psychodynamic and dialectical behavior therapies for borderline personality disorder. *Personality Disorders, 9,* 437–446. http://dx.doi.org/10.1037/per0000276

Sharp, C., Wright, A. G., Fowler, J. C., Frueh, B. C., Allen, J. G., Oldham, J., & Clark, L. A. (2015). The structure of personality pathology: Both general ('g') and specific ('s') factors? *Journal of Abnormal Psychology, 124,* 387–398. http://dx.doi.org/10.1037/abn0000033

Shaver, P. R., & Mikulincer, M. (2005). Attachment theory and research: Resurrection of the psychodynamic approach to personality. *Journal of Research in Personality, 39,* 22–45. http://dx.doi.org/10.1016/j.jrp.2004.09.002

Sidik, K., & Jonkman, J. N. (2002). A simple confidence interval for meta-analysis. *Statistics in Medicine, 21,* 3153–3159. http://dx.doi.org/10.1002/sim.1262

Skodol, A. E., Grilo, C. M., Pagano, M. E., Bender, D. S., Gunderson, J. G., Shea, M. T., . . . McGlashan, T. H. (2005). Effects of personality disorders on functioning and well-being in major depressive disorder. *Journal of Psychiatric Practice, 11,* 363–368. http://dx.doi.org/10.1097/00131746-200511000-00002

Stern, B. L., Caligor, E., Clarkin, J. F., Critchfield, K. L., Horz, S., MacCornack, V., . . . Kernberg, O. F. (2010). Structured Interview of Personality Organization (STIPO): Preliminary psychometrics in a clinical sample. *Journal of Personality Assessment, 92,* 35–44. http://dx.doi.org/10.1080/00223890903379308

Strickland, C. M., Hopwood, C. J., Bornovalova, M. A., Rojas, E. C., Krueger, R. F., & Patrick, C. J. (2019). Categorical and dimensional conceptions of personality pathology in *DSM–5*: Toward a model-based synthesis. *Journal of Personality Disorders, 33,* 185–213. http://dx.doi.org/10.1521/pedi_2018_32_339

*Svartberg, M., Stiles, T. C., & Seltzer, M. H. (2004). Randomized, controlled trial of the effectiveness of short-term dynamic psychotherapy and cognitive therapy for cluster C personality disorders. *The American Journal of Psychiatry, 161,* 810–817. http://dx.doi.org/10.1176/appi.ajp.161.5.810

Viechtbauer, W. (2010). Conducting meta-analyses in R with the meta for package. *Journal of Statistical Software, 36,* 48. http://dx.doi.org/10.18637/jss.v036.i03

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: Empiricially, "all must have prizes. *Psychological Bulletin, 122,* 203–215. http://dx.doi.org/10.1037/0033-2909.122.3.203

Weissman, M. M., & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. *Archives of General Psychiatry, 33,* 1111–1115. http://dx.doi.org/10.1001/archpsyc.1976.01770090101010

*Winston, A., Laikin, M., Pollack, J., Samstag, L. W., McCullough, L., & Muran, J. C. (1994). Short-term psychotherapy of personality disorders. *The American Journal of Psychiatry, 151,* 190–194. http://dx.doi.org/10.1176/ajp.151.2.190

Wright, A. G. C., Hopwood, C. J., Skodol, A. E., & Morey, L. C. (2016). Longitudinal validation of general and specific structural features of personality pathology. *Journal of Abnormal Psychology, 125,* 1120–1134. http://dx.doi.org/10.1037/abn0000165

Yeomans, F. E., Clarkin, J. F., & Kernberg, O. F. (2015). *Transference-focused psychotherapy for borderline personality disorder*. Washington, DC: American Psychiatric Publishing.

Young, J. E., Klosko, J. S., & Weishaar, M. (2003). *Schema therapy: A practitioner's guide*. New York, NY: Guilford Press.

Zickgraf, H. F., Chambless, D. L., McCarthy, K. S., Gallop, R., Sharpless, B. A., Milrod, B. L., & Barber, J. P. (2016). Interpersonal factors are associated with lower therapist adherence in cognitive-behavioural therapy for panic disorder. *Clinical Psychology and Psychotherapy, 23,* 272–284. http://dx.doi.org/10.1002/cpp.1955