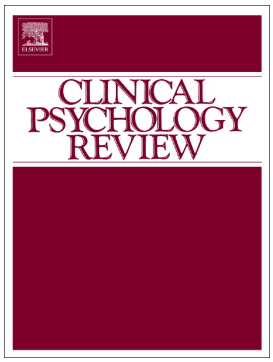


Brief psychological interventions for borderline personality disorder. A systematic review and meta-analysis of randomised controlled trials

A.J. Spong, I.C.H. Clare, J. Galante, M.J. Crawford, P.B. Jones

PII: S0272-7358(20)30125-2

DOI: https://doi.org/10.1016/j.cpr.2020.101937

Reference: CPR 101937

To appear in: *Clinical Psychology Review*

Received date: 8 June 2020

Revised date: 7 September 2020

Accepted date: 13 October 2020

Please cite this article as: A.J. Spong, I.C.H. Clare, J. Galante, et al., Brief psychological interventions for borderline personality disorder. A systematic review and meta-analysis of randomised controlled trials, *Clinical Psychology Review* (2020), https://doi.org/10.1016/j.cpr.2020.101937

Brief Psychological Interventions for Borderline Personality Disorder. A systematic review and meta-analysis of randomised controlled trials

Spong, A.J.[a], Clare, I.C.H.[b], Galante, J.[c], Crawford, M.J.[d], Jones, P.B.[e]

Background: People with Borderline Personality Disorder (BPD) have limited access to long term psychological therapies. Briefer interventions have been developed but trial evidence to support their use has not been reviewed.

Aims: To examine whether psychological interventions for adults with BPD of six months duration or less improve symptoms, mood, self-harm, suicidal behaviour, and service use.

Methods: The protocol was prospectively registered (PROSPERO CRD42017063777). Database searches were conducted up to April 2020. Inclusion, data extraction and risk of bias were assessed in duplicate. We identified 27 randomised controlled trials. We conducted random-effects meta-analyses sub-grouping data into delivery method, additional support, and comparison type.

Results: High levels of bias were found for attrition and reporting. Heterogeneity was high in some pooled data. Borderline symptom reductions were greatest for interventions including additional support (SMD. -1.23, 95% C.I. -2.13, -0.33). Planned generic support may be as effective as specialist interventions for borderline symptoms (SMD = -0.11, 95% C.I. -0.51, 0.29) and social functioning (SMD= -0.16., 95% C.I. -0.65, 0.33). Follow-up was limited and direct comparison with post-intervention results was unreliable.

Conclusions: Short-term interventions may be effective. Access to additional support has an impact on outcomes. It is unclear if symptomatic change is sustained.

[a] Amanda J. Spong (corresponding author) D.Clin.Psy,. Principal Clinical Psychologist, Cambridgeshire & Peterborough NHS Foundation Trust (CPFT), United Kingdom. Tel: 07710242720. Email: amandaspong68@gmail.com

Present address: East London Foundation Trust, The Lawns Resource Centre, Biggleswade, SG18 0PT

[b] Isabel C.H. Clare PhD. Consultant Clinical & Forensic Psychologist, 1. National Institute of Health Research (NIHR) Applied Research Collaboration (ARC) East of England, 2. Department of Psychiatry, University of Cambridge, Cambridge United Kingdom; 3. CPFT

[c] Julieta Galante PhD., NIHR Postdoctoral Fellow, Department of Psychiatry, University of Cambridge, United Kingdom; second affiliation: NIHR ARC East of England

[d] Mike J. Crawford MD, Professor of Mental Health Research, Division of Psychiatry, Imperial College London, United Kingdom.

[e] Peter B. Jones PhD, Professor of Psychiatry, University of Cambridge, Herchel Smith Building, Addenbrooke's Hospital, Cambridge, United Kingdom.

Keywords

Borderline personality disorder

Psychological interventions

Systematic review

Short-term treatment

Introduction

Borderline personality disorder (BPD) is a severe mental health condition that affects 1-2% of the population (Coid, Yang, & Tyrer, 2006) and is characterised by significant difficulties in emotion regulation, the ability to develop and sustain relationships, and the ability to regulate impulsive behaviour, including self-harm and suicidal behaviour (American Psychological Association, 2013). A range of long-term psychological therapies have been developed that have been shown to improve the mental health of people with borderline personality disorder (Omar, Tejerina-Arreal, & Crawford, 2014; Stoffers-Winterling, Völlm, Rücker, Timmer, & Lieb, 2012), and are included in recommendations and guidelines for the treatment of people with BPD (American Psychiatric Association, 2001; National Institute for Health and Care Excellence, 2009; National Health and Medical Research Council, 2012; Simonsen et al., 2019). Long term psychological therapies have limited availability (Paris, 2013) and many people with severe borderline personality disorder are unable to engage or drop out of treatment before it is completed (Crawford et al., 2009; McMurran, Huband, & Overton, 2010). There has been a reluctance to deliver brief psychological therapies due to concerns that they may be unhelpful or even harmful for people who can find it difficult to form trusting relationships and cope when relationships end (National Institute for Health and Care Excellence, 2009). However, in response to increased recognition of the needs of people with BPD and increased demand for treatment, efforts have been made to develop and test the effectiveness of shorter-term interventions. Stoffers et al. (2012) identified 12 studies of interventions of 6 months duration or less. Subsequent reviews have mainly identified small-scale feasibility studies of shorter-term interventions with inconclusive results (Cristea et al., 2017).

Evidence of effectiveness of psychological interventions for BPD and shorter, manualised treatments is fundamental to increasing access. To date, there has been no comprehensive review of brief psychological interventions for BPD to help inform developments in this field; this review aims to address this shortfall. We set out to examine whether brief psychological interventions, lasting six months or less, are effective for adults with BPD. We also examined whether the form and content of interventions was associated with their effectiveness.

Method

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines are used to report this review. Ethics and consent statements are not reported as this is a review of previously published studies. Details of the protocol were prospectively registered on PROSPERO (www.crd.york.ac.uk; CRD42017063777), setting out the parameters and methodology to be used. The decision to analyse stand-alone interventions against those delivered as an add-on to other treatment was taken after the publication of the protocol as subsequent evidence from Cristea et al. (2017) suggested that the effect of additional support remains unclear.

Search strategy

Studies were identified through searches of six bibliographical databases (PubMed, PsychINFO, Embase, Medline, CINAHL, and the Cochrane Central Register of Controlled Trials) from 1980. Searches were carried out on 5th April 2017, 1st October 2018, and 31st March 2020, and were restricted to English language publications. Search terms were based on those used by NICE in the production of guidance for BPD. These include terms related to psychological therapies ('psychotherapy*' or 'intervention*'), borderline personality disorder ('borderline$' and 'personalit$'), and clinical trials. These were augmented with terms that would describe brief or low

intensity psychological interventions, such as skills training and psycho-education ('brief*' or 'psychoeducation*'). The reference lists of other systematic reviews of the treatment of borderline personality disorder were checked, and the researchers of a National Institute for Health Research funded project developing a brief intervention for borderline personality disorder were contacted to check for additional papers.

Study selection

Population: Studies were included if they focused on adults aged 18 years or over, with a primary diagnosis of borderline personality disorder as defined either by a classification system, a validated structured clinical interview, or by a psychiatrist. Studies that identified multiple subtypes of personality disorder were included, but those that focused exclusively on other sub-types of personality disorder, e.g. antisocial personality disorder, were excluded. The reason is that people with at least moderate levels of personality disorder pathology often meet diagnostic criteria for more than one sub-type, and this usually includes borderline personality disorder. Studies that involved participants with co-morbid conditions, such as substance use or PTSD, were excluded if the intervention and outcomes were focused on the reduction of those symptoms, as opposed to symptoms of borderline personality disorder.

Intervention: We included studies that tested interventions based on a psychological theory or model, including psychotherapy, psychological therapy, psycho-education, or skills training, delivered either individually or in groups. Interventions of 6 months duration or less were included. This time frame was chosen because there are a number of treatment protocols in use which are of six months or less duration e.g. STEPPS, DBT skills training, and a previous Cochrane review (Soffers et al. 2012) identified a small cluster of studies of interventions of six months or less duration. Consistent with a previous high-quality review (Cristea et al. 2017) studies that included medication as part of one or more treatment conditions were excluded, but those in which medication was included in ongoing treatment as usual were not.

Comparators: Only randomised clinical trials were included. Studies comparing two or more interventions for BPD were only included if they also used a control group. The definition of a control group was broad and included other psychological treatments not specifically developed for the treatment of personality disorder. These could include treatment as usual (TAU), structured clinical management (SCM), or a non-manualised or generic psychotherapy. In line with the review by Cristea et al. (2017), comparisons between components of the same psychotherapy, e.g. emotion regulation skills training and interpersonal effectiveness training in DBT, were excluded because of expectations of similar efficacy.

Outcomes: Primary outcomes were in the following four categories; core symptoms of borderline personality disorder, mood, self-harm and suicidal behaviour, and service use. Secondary outcomes were social functioning and general mental health.

Selection process

Titles and abstracts were imported to a reference management system and screened for eligibility and duplicates by A.S. and I.C.H.C.. When a trial included two interventions that were eligible for this review, they were combined for analysis. Full texts of studies meeting inclusion criteria were then screened for the duration of the intervention condition. Reasons for excluding studies are summarised in the PRISMA flow diagram (Figure 1). Each stage of the review was carried out independently by the researchers, and decisions were recorded on separate spreadsheets.

Data extraction and quality assessment

Data were extracted by A.S. and I.C.H.C. independently using separate spreadsheets. Quality was assessed using the domains of the Cochrane Risk of Bias tool (Higgins et al., 2011), which assesses sources of bias in randomised control trials. The tool assesses the likelihood of bias in selection, detection, attrition and reporting. Studies were classified as having high, low or unclear risk of bias in each category and disagreements were resolved through discussion. For the purposes of this review attrition was defined as when a participant consented to treatment but did not complete all of outcome measure questionnaires over the course of the study. Studies were classified as having high levels of attrition if the average drop-out rate exceeded 20% of the total sample. If reasons for drop out were specified in the paper then the drop out threshold was increased to 30% of the sample.

In addition, information was collected on therapy type, duration in weeks and number of sessions for intervention and control arms, mode of delivery, participant characteristics, whether the intervention was stand-alone or adjunctive to other treatment, the inclusion of individual sessions, the type of control condition, the involvement of the treatment developer in the study team, and the presence of therapist supervision and monitoring of adherence to treatment protocols. The type of control condition was classified as either manualised or non-manualised. A manualised control condition was defined as a non-specialist intervention delivered in a planned manner, in accordance to a treatment protocol or manual. A non-manualised control condition usually included waiting list or treatment as usual that was not delivered in a planned way.

[Insert Figure 1. PRISMA flow diagram]

Analysis

Mean and standard deviation data for experimental and control conditions for each outcome category were entered into a computer program (Review Manager,2014) using random effects analysis to calculate standard effect sizes for the following subgroups;

- Any eligible interventions versus manualised and non-manualised control conditions
- Interventions subdivided as using group only, individual only, and mixed individual and group sessions versus any control group
- Interventions subdivided as stand-alone or adjunctive interventions versus any control group
- Interventions subdivided depending on intervention name/type versus any control group

Post-intervention data were calculated up to 8 weeks following the end of treatment and follow-up data was included up to 2 years after the end of treatment. This cut-off is consistent with that used by Cristea et al. (2017), who surmised that a time span beyond 2 years increased the risk of biasing effects from factors extraneous to the intervention. The same approach was taken for primary and secondary outcomes. If papers did not include the necessary data the authors were contacted. If the authors failed to respond after two months the study was excluded from the analysis. Where possible intention to treat (ITT) data were used. If a study reported on multiple outcomes in the same category, the most commonly used was selected, so that each trial reported on just one set of data in each outcome category for each time point. Effect sizes for continuous outcomes (Hedges' *g*) were computed, using a random effects model, to calculate standard mean difference between intervention and control group data for each outcome category. Where dichotomous outcomes were reported, the number of observed events was used to calculate odds ratios. Heterogeneity was assessed with the I^2 statistic. 0% indicates no heterogeneity, with thresholds at 25%, 50% and 75% indicating thresholds for low, moderate and high. These thresholds are consistent with an earlier published review (Cristea et al., 2017).

Results

The review identified 28 papers, reporting on 27 different studies. Two papers by Bos were published using data from the same study; one reporting on participants who met diagnostic criteria for BPD (Bos, van Wel, & Appelo, 2011) and the other reporting on a larger sample which included participants with sub-threshold symptoms (Bos, van Wel, Appelo & Verbraak 2010). The paper that only included participants who met diagnostic criteria was selected. Another study (Andreoli, Burnand, & Cochennec, 2016) evaluated the delivery of the same brief intervention in two intervention arms, compared to a control; one was delivery by nurses, the other was delivery by psychologists. There was no significant difference between the two treatment arms and the data were combined for analysis. Included studies are marked by * in the References section.

Study characteristics

Within the included group of 27 papers, twenty-three studies were based in community or outpatient settings, with only two studies in in-patient settings. The majority of studies diagnosed participants using a standardised means of assessment. Four studies (15%) diagnosed participants from a structured interview by a psychiatrist using DSM-IV criteria. There were a total of 2219 participants; 1164 in treatment conditions and 1055 in control conditions. The length of intervention ranged from 2 to 24 weeks, with a mean intervention length of 13.9 weeks. The most common intensity of delivery was weekly sessions, but there was some variation across studies. The total number of hours of each intervention was not specified in every study. Where the data were given, it ranged from 7.5 to 60 hours, with a mean of 27.2 hours. The majority of studies were conducted either in Europe or North America, with just four studies coming from outside these areas: two from Australia, one from Iran and one from Taiwan.

There were 15 different brief psychological interventions evaluated (a brief description of interventions is attached to Supplementary Figure 1 in Appendix). DBT skills groups were the most frequently studied (5 studies), but there were generally low levels of replication. Fifteen studies (56%) evaluated interventions delivered in conjunction to treatment as usual, whilst the remaining 12 studies examined stand-alone treatments. Thirteen studies (48%) were of group only sessions, seven (26%) of individual only sessions, and seven (26%) included mixed group and individual. The majority of studies used a non-manualised control condition, normally treatment as usual, with only 7 out of the 25 studies (26%) employing a manualised control condition. These included either supportive groups or therapy not specialised to the treatment of borderline personality disorder, and one paper used general psychiatric management (GPM).

In 11 studies (41%), the treatment developer was involved in delivering treatment, or supervising those delivering treatment or the research team. In a further two studies there was insufficient information to judge whether the treatment developer was involved. Fourteen studies (52%) were carried out independently of the treatment developer. Eighteen studies (67%) did not provide information on the amount of face to face contact received by participants in the treatment and control conditions. Of the remaining nine studies which reported these data, seven reported equal levels of contact in both treatment and control conditions. Seventeen studies (63%) did not provide information about whether there was any form of monitoring of how treatment protocols were delivered. This could be in the form of supervision of therapists or videotaping of treatment sessions. Without this, it was not possible to ensure that patients received the treatment as it was designed.

Risk of bias

Levels of risk of bias varied across the studies. The categories with highest levels of risk were attrition and reporting bias. Only 9 of the 27 studies (33%) reported publishing a protocol for the study prior to the commencement of the trial. This means that there was the potential for bias in the publication of results, with only favourable data being reported. Twelve of the studies (44%) reported high levels of attrition across both intervention and control groups.

The risk of bias assessments and study characteristics mean that it is possible that some treatment effects could be influenced by selective reporting of results, by 'dose effects' rather than the content of interventions, or by variation in the delivery of interventions. In addition, although most studies used intention to treat methods in their analysis, around half of the studies reported on interventions that a significant number of participants were not able to complete. Notwithstanding these comments, low risk of bias was found in the majority of studies in four of the categories (see Supplementary Figures 2, 3 & 4 for full risk of bias summary).

Main effects post-intervention

i) Are interventions more effective than manual sed or non-manualised controls?

Table 1: Post-test effect sizes by control group

	Non-manualised control			**Manualised control**			p Value
Variable	**No. of trials (no. of participants)**	**Hedges' g (95% CI)**	**I^2%**	**No. of trials (no. of participants)**	**Hedges' g (95% CI)**	**I^2%**	
Borderline symptoms	7 (576)	-1.24 (-2.12, -0.37)	95	4 (266)	-0.11 (-0.51, 0.29)	61	0.02
Depression	9 (628)	-0.94 (-1.30, -0.58)	73	3 (192)	-0.45 (-0.78, -0.12)	19	0.05
Anxiety	4 (144)	-1.14 (-1.68, -0.60)	52	2 (110)	-0.39 (-0.77, -0.01)	0	0.03
Social functioning	8 (871)	-0.82 (-1.21, -0.44)	85	2 (154)	-0.16 (-0.65, 0.33)	58	0.04

Studies using a non-manualised control condition (n=20) showed large to very large effect sizes across all outcomes. The difference between these effect sizes and those for studies using a manualised control condition was statistically significant across all pooled outcome data. The largest effect size was for borderline symptoms but the level of heterogeneity was also large. Only the outcome of anxiety had more acceptable levels of heterogeneity.

Only seven studies used a manualised control and effect sizes were significantly smaller than for non-manualised control conditions. The effect size for borderline symptoms and social functioning

showed little difference between the intervention and control groups. The largest effect size was for depression, with low levels of heterogeneity in the data.

The very high levels of variance in data suggested that additional factors may have been influencing outcomes. Based on the conclusions of another review (Stoffers et al., 2012) post-hoc analyses were used to examine the effects of the involvement of the treatment developer. For the outcome of borderline symptoms, five studies had the involvement of the treatment developer. Pooled data showed a very large effect size (-1.55) but high levels of heterogeneity (I^2**%** = 96%). Four studies did not have the involvement of the treatment developer. This group showed a moderate effect size (-0.50) and no heterogeneity in the pooled data.

ii) *Does the mode of delivery determine treatment effectiveness?*

Table 2: Post-test effect sizes by session type

Variable	**Individ sessions only**			**Group sessions only**			**Group + individ sessions**			p value
	No. of trials (No. of Ps)	**Hedges' g (95% CI)**	**I^2%**	**No. of trials (No. of Ps)**	**Hedges' g (95% CI)**	**I^2% %**	**No. of trials (No. of Ps)**	**Hedges' g (95% CI)**	**I^2%**	
Borderline symptoms	N/A			7 (440)	-1.21 (-2.26, -0.17)	95	3 (328)	-0.35 (-0.71, -0.02)	63	0.12
Self-harm	N/A			2 (83)	-0.84 (-1.29, -0.39)	0	3 (145)	-0.24 (-0.57, 0.08)	0	0.04
Depression	N/A				-0.89 (-1.29, -0.49)	78	2 (104)	-0.61 (-1.01, -0.22)	0	0.34
Social functioning	3 (192)	-0.42 (-0.28, 0.14)	73	3 (265)	-1.12 (-2.18, -0.07)	93	3 (536)	-0.49 (-0.67, -0.32)	0	0.49
General mental health	3 (193)	-0.33 (-0.95, 0.29)	78	4 (318)	-0.57 (-2.04, 0.89)	97	N/A			0.76

There were more studies using group-only sessions across all outcome categories (n=13), compared to those using individual only (n=7) or mixed sessions (n=7). The effect sizes for group-only interventions were large across most of the outcome categories. However, heterogeneity was very high for all outcomes except for self-harm and this raises questions about the reliability of the pooled data. The only statistically significant result was found comparing group-only interventions with mixed delivery interventions for the outcome of self-harm, where group-only interventions had a larger effect size.

Studies using only individual sessions found moderate effect sizes for general mental health and social functioning. These effects were smaller compared to studies using group only or mixed sessions, although this difference did not reach statistical significance.

The effect sizes for studies using mixed sessions were small to moderate, and the heterogeneity levels were negligible in all but one outcome category. Again, the largest effect size was for the outcome of depression.

Due to the high levels of variance in data for group-only studies we carried out post-hoc analysis. For the outcome of borderline symptoms, four studies had the involvement of the treatment developer and showed a large effect size (Hedges' g = -1.95) although very high levels of heterogeneity (I^2**%**= 92%). Two further studies did not have the involvement of the treatment developer. This sub-group had a low effect size (-0.26) but high heterogeneity (I^2**%** =85%). For the outcome of depression, five studies using group-only interventions had the involvement of the treatment developer. Pooled data showed a very large effect size (-1.16) but high heterogeneity (I^2**%**=81%). A further three studies did not have the involvement of the treatment developer. These data showed a smaller, but still large, treatment effect (-0.72) and moderate heterogeneity (56%).

iii) Does delivery of interventions as stand-alone or adjunctive to other treatment impact outcomes?

Table 3. Post-test effect sizes by additional support

Variable	Stand-alone			Adjunctive			p value
	No. of trials (No, of Ps)	Hedges' g (95% CI)	I^2%	No. of trials (No, of Ps)	Hedges' g (95% CI)	I^2%	
Borderline symptoms	4 (323)	-0.13 (-0.50, 0.24)	92	7 (519)	-1.23 (-2.13, -0.33)	95	0.03
Self-harm	2 (61)	-0.21 (-0.72, 0.29)	0	4 (197)	-0.62 (-0.96, -0.28)	24	0.19
Depression	5 (382)	-0.51 (-0.74, -0.29)	0	7 (438)	-0.99 (-1.44, -0.55)	77	0.06
Anxiety	3 (159)	-0.59 (-1.03, -0.15)	39	3 (124)	-1.12 (-1.82, -0.41)	68	0.22
Social functioning	3 (298)	-0.48 (-0.71, -0.25)	0	6 (695)	-0.78 (-1.34, -0.22)	91	0.34
General mental health	3 (301)	-0.41 (-1.70, 0.89)	95	6 (490)	-0.30 (-1.18, 0.58)	95	0.90

There were more trials using interventions in addition to other support (n=15), normally treatment as usual, than trials of stand-alone interventions (n=12). Trials of stand-alone interventions were not much more effective than controls for borderline symptoms. Effect sizes were higher for adjunctive interventions than stand-alone interventions, with the exception of general mental health, ranging

from moderate to very large. However, heterogeneity levels were also generally quite high and the only statistically significant result was for the outcome of borderline symptoms.

iv) Is any intervention more effective than others?

Table 4: Post-test effect sizes by intervention

	DBT Skills			DBT-SF			ER Group			STEPPS			p value
Var	**Trials (Ps)**	**Hedges' g**	**I^2%**	**Trials (Ps)**	**Hedges' g**	**I2 %**	**Trials (Ps)**	**Hedges' g**	**I^2%**	**Trials (Ps)**	**Hedges' g**	**I^2%**	
BPD	3 (325)	-0.40 (-0.67, -0.13)	88	N/A			2 (83)	-1.51 (-2.35, -0.68)	58	2 (237)	-1.95 (-5.13,1.23)	99	0.00001
Dep	3 (225)	-0.57 (-0.93, 0.20)	45	N/A			2 (83)	-1.38 (-1.87, -0.9)	0	N/A			0.08
SH	N/A			2 (93)	-0.21 (-0.61, 0.2)	0	2 (83)	-0.8 (-1.29, -0.39)	0	N/A			0.04

Only four interventions had sufficient replication across outcome categories to pool data and compare effect sizes post-test. For BPD symptoms, Emotion Regulation group (ER group) had greater post-intervention effect sizes than DBT skills, and lower levels of heterogeneity. However, STEPPS had a statistically significantly greater effect size than either DBT skills or ER group, but the confidence intervals and heterogeneity of the data for STEPPS were so large that the analysis was considered uninformative. ER group had a statistically significantly greater effect size than DBT skills for the outcome of depression, and greater than DBT SF (DBT short form) for self-harm.

Main effects at follow-up

Fourteen studies included follow-up periods in their design, but not all of these published their follow-up data or were able to provide it to the researchers.

v) Does the mode of delivery determine treatment effectiveness on follow-up?

Table 5: Follow-up effect sizes by intervention type

	Individual sessions only			Group only sessions			Group + individual sessions			p
Variable	**No. of trials (no. of Ps)**	**Hedges' g (95% CI)**	**I^2%**	**No. of trials (no. of Ps)**	**Hedges' g (95% CI)**	**I^2%**	**No. of trials (no. of Ps)**	**Hedges' g (95% CI)**	**I^2%**	
Borderline symptoms	N/A			2 (159)	-1.26 (-2.94, 0.41)	96	3 (312)	-0.11 (-0.33, 0.11)	0	0.18
Depression	N/A			3 (283)	-0.12 (-0.49,	57	2 (192)	0.01 (-0.29,	0	0.63

					0.24)			0.27)		
Social functioning	N/A			2 (201)	-1.24 (-3.13, 0.64)	97	2 (390)	-0.59 (-0.82, -0.37)	11	0.50

On follow-up there was no difference between the intervention and control group data for mixed session designs across BPD symptoms and depression. However, there was a moderate effect size for social functioning. Studies using group only interventions showed very large effects for borderline symptoms and social functioning on follow-up. However, although this is a smaller group of studies than for the post-test data, the heterogeneity in the follow-up data remains extremely high. There were no statistically significant differences found between the subgroups.

vi) Are interventions more effective than non-manualised or manualised controls on follow-up?

Table 6: Follow-up effect sizes by control group

	Non-manualised control			**Manualised control**			p value
Variable	**No. of trials (no. of participants)**	**Hedges' g (95% CI)**	**I^2%**	**No. of trials (no. of participants)**	**Hedges' g (95% CI)**	**I^2%**	
Borderline symptoms	2 (204)	-0.22 (-0.50, 0.06)	0	3 (267)	-0.80 (-2.03, 0.42)	95	0.36
Depression	2 (208)	0.07 (-0.20, 0.34)	0	3 (267)	-0.17 (-0.44, 0.11)	23	0.23

On follow-up, the effect sizes were negligible or small for all outcomes, with the exception of BPD symptoms for studies with manualised controls. However, the heterogeneity of the data in this group was extremely high. There were no statistically significant differences found between the subgroups.

Discussion

Using search terms adapted for brief psychological interventions, we identified more than twice as many studies than in previous reviews (Omar et al., 2014; Stoffers et al., 2012). Although the term 'brief' was defined as up to six months, the average length of intervention was just over three months. Despite the recommendations from NICE (National Institute for Health and Care Excellence, 2009) that treatment should be no less than three months' duration, the literature contains many evaluations of interventions that last around three months. The results of this review may have important implications for service development and clinical practice.

Pooling and analysis of data on the outcome of self-harm were restricted because of the variation across studies in the measurement of this variable. Some studies measured suicide attempts (Lin, Huei-Chen, & Yung-Wei, 2018), while others measured suicide and other self-harm behaviours (McMain, Guimond, & Barnhart, 2017). Different time frames were also used, for example, behaviours that occurred within the last three months (Carter, Willcox, & Lewin, 2010) or over the lifetime (McMain et al., 2017). While some studies used continuous data (Cater et al., 2010; Evans, Tyrer, & Catalan, 1999; Koons, Robins, & Tweed, 2001; Weinberg, Gunderson, & Hennen, 2006), others measured percentages or proportion of participants reporting self-harming behaviours (Andreasson, Krogh, & Wenneberg, 2016; Lin et al., 2018,).

For similar reasons, it was not possible to pool and analyse any data on service use. Eleven studies included a measure of service use but there was very little consistency in its definition and reporting. The data were mostly nominal, counting the number of contacts or admissions to hospital, but there were also examples of the use of contact with any mental health professional (Huband, McMurran, & Evans, 2007), visits to hospital emergency departments (Andreoli et al., 2016; Evans et al., 1999; Soler, Pascual, & Tiana (2009) and contact with community crisis teams (Blum, St. John, & Pfohl 2008; McMurran, Crawford, & Reilly, 2016). In addition to the variation in definitions of service use, there was a lack of consistency in the unit of measurements, with some studies measuring frequency of contact (Clarke, Thomas, & James, 2013) and others the length of admission (Koons et al., 2001).

Impact of comparator

Regardless of intervention type, the comparison condition had a significant impact on the size of treatment effect. Comparing treatments either to waiting list or treatment as usual showed moderate to large effects in favour of the intervention, although levels of heterogeneity for all outcomes was moderate to high. However, when interventions were compared to planned, regular and structured contact there were small or negligible effects of the intervention and these differences were statistically significant across all outcomes. Heterogeneity was generally lower in this grouping, suggesting the data were more reliable. This raises the question of what benefits patients: the specialist nature of the intervention or something more generic such as planned and predictable contact with a health care professional? Addressing this issue could have major implications for the staffing and costs of services but needs further examination. Previous research comparing longer-term psychological therapies with generalist, structured interventions, such as Structured Clinical Management (SCM) or General Psychiatric Management (GPM) has found these manualised control conditions to be as effective as specialist therapies (Bateman & Fonagy, 2009; McMain et al. 2009; Clarkin, Levy, Lenzenweger, & Kernberg, 2007) but only one study in this review used GPM as a control (Kramer, Kolly, & Berthoud, 2004). There needs to be further comparison of brief psychological interventions with manualised control conditions.

Impact of mode of delivery

Interventions delivered through individual sessions had a moderate effect sizes for social functioning and general mental health. However, there was not enough replication to pool data for primary outcomes. Only two studies collected data on borderline symptoms (Clarke et al. 2013; Kramer et al. 2004) and study differences meant that the data could not be pooled. There is not therefore enough evidence to comment on whether individual interventions are effective for BPD symptoms, mood or self-harm.

Across all outcome categories, studies using group only interventions showed larger effects compared to mixed intervention studies. However, with the exception of self-harm, the heterogeneity of the pooled data for the group-only studies was moderate to high, making it difficult to draw conclusions from this data. This group included studies of eight different interventions and it may be that each intervention had an effect on an aspect of participants' difficulties.

Blum et al. (2008) showed very strong results in favour of STEPPS plus treatment as usual, a group-only intervention, but visual inspection of forest plots showed the data from this study as an outlier compared to other pooled data. There were a number of sources of bias in the study, including the involvement of the treatment developer, large differences in levels of contact between the treatment and control conditions, and reporting data only from participants completing the intervention. The study by Blum was not the only one of the group-only intervention studies in which the treatment developer was involved in the evaluation. Indeed, they were included in half of such studies. The analysis suggested that this had a strong positive impact on effect sizes. In studies without the involvement of the treatment developer, effect sizes were reduced by half.

The only statistically significant difference between effect sizes was between group-only and mixed session interventions for the outcome of self-harm, in favour of group-only interventions. It may be that the social context of interventions impacts the motivation to change self-harming behaviours in particular. However, given the possible confounding factors in the group-only sub-group, the high levels of heterogeneity in the pooled data, and the dearth of pooled data for individual only interventions, it is not possible to conclude whether mode of delivery has a broader impact on effect sizes across different outcomes. Mode of delivery can have a significant implication on resources needed to provide interventions and it would seem important to be able to identify whether this is a relevant factor. Further research is needed that compares mode of delivery and uses consistent outcome measures to enable pooling of data.

Impact of additional support

The other factor that had a strong influence on treatment effects was the presence or absence of ongoing support to the delivery of the intervention. Ongoing support had an even stronger effect than the type of control group. With the exception of the outcome of general mental health, the heterogeneity levels were lower across all outcomes for stand-alone interventions. This finding may have reflected problems in defining general mental health and it was possible that the outcome measures used in different studies were assessing different facets of mental health and mental wellbeing. It appears that delivering brief psychological interventions, whether group based or not, in isolation from other forms of support resulted in small to moderate effect sizes. Across the majority of outcomes, studies that used interventions as an adjunct to ongoing support showed effect sizes that were at least double those in which such support was absent, regardless of the intervention delivery or control condition. This difference reached statistical significance for the outcome of borderline symptoms. McMurran et al. (2016) was one study in the adjunctive sub-group. This was a large and well-structured RCT, using mixed individual and group sessions, with generally low levels of bias. This study conducted semi-structured interviews with all participants after treatment and at 72-week follow-up, and qualitatively analysed their responses. Many participants found the content of the groups to be relevant to their difficulties and easy to engage with, but some reported the group format triggered vulnerabilities and unfavourable comparisons with others. The authors conclude that "service users and providers have also highlighted the importance of co-ordinating and integrating psychological treatments within teams that can provide additional support for patients at times of crisis." (p.86). "The null findings of the PEPS trial suggest

that the intervention may have been too brief to have had an effect, at least as a stand-alone treatment divorced from good clinical care." (p.85).

Comparative efficacy of interventions

There are a number of reasons why the results of the meta-analyses should be treated cautiously. As noted, not all interventions or outcome categories involved sufficient replication to allow results to be pooled, and those that did contained a small number of trials. In addition, the heterogeneity of BPD data for DBT skills were very high while the levels for the STEPPS trials was so high that it makes the data unreliable. The ER group data appear superior, but contained a number of biases. Although there were low levels of drop out, both ER group trials involved the treatment developer (Gratz). The interventions were adjunctive to treatment as usual, which included an individual therapist who supported the group intervention. Two of the three DBT skills studies used stand-alone treatments (Lin et al. 2018; Soler et al. 2009) so there was an important difference in the design of the studies that made direct comparison of these interventions problematic.

Stability of change on follow-up

It is difficult to make any direct comparison of post-test effect sizes and those found in pooled data on follow-up. The studies in almost every outcome category were not consistent. However, what is striking is that mixed delivery studies showed low or negligible effects at follow-up for BPD symptoms and depression and that the moderate effect sizes post-test have almost disappeared. Although effect sizes for BPD symptoms and social function are very high for group only studies, the levels of heterogeneity of the data make it difficult to draw conclusions. The effect sizes for control conditions were more consistent, with the exception of BPD symptoms for studies with manualised control conditions. However, again the levels of heterogeneity were very high. When heterogeneity was at an acceptable level, most follow-up treatment effects disappeared.

Limitations

This review adheres to the published protocol and extends the number of studies of brief psychological interventions for BPD identified in previous reviews. However, there are some limitations.

There were three papers that could not be included as the text, apart from the abstract, was not in English. The decision to exclude non-English papers was made *a priori* and specified in the published protocol because of the lack of availability of translation services for such technical documents. It is acknowledged that exclusion on this basis can contribute to publication bias. One paper reported on a feasibility study of an adolescent population and as such would not have met inclusion criteria. Another reports on a study of STEPPS plus individual therapy and may include the same participants as reported in the Bos 2011 study, included in this review. The third is a controlled trial of a brief analytic psychotherapy and may have met inclusion criteria, but it is not possible to draw a conclusion in the absence of access to the full paper. Therefore, one of the omitted papers may have been included in the review if it had been possible to translate. While this does introduce publication bias the impact on the pattern of results may not have changed the overall conclusions of the review.

The difficulty in pooling data on levels of service and self-harm are important. Self-harm behaviours cause distress to service users and correlate highly with completed suicide. More consistent definitions and measures need to be used so that the effectiveness of interventions can be more

readily evaluated. The impact of interventions on service use, especially crisis services such as emergency department and inpatient admissions, is important in evaluating their cost and impact on services. Some outcome categories included data from only a small number of studies. In addition, many of the studies in the review used small sample sizes and it is possible that some of the analyses were underpowered and therefore effect sizes may not have been reliable. Participants in many studies were predominantly women and seven studies used female only samples.

Creating three different categories of sub-groups for analysis (control condition, mode of delivery, and presence of additional support) may have introduced confounding variables to the sub-group analysis, influencing the findings. There were insufficient studies to be able to control for the impact each variable had on the other, and this is a limitation of the statistical power of the analyses.

Many of the outcome measures were self-reported and although detection bias was good across most studies the data may have been influenced by response bias. The use of medication, particularly in treatment as usual conditions, was neither standardised or consistently reported and so could have confounded therapy effects.

Finally, in much of the pooled data, where effect sizes were high, levels of heterogeneity were also high. A number of factors may have influenced this finding but they need to be taken into account in drawing conclusions about the influence of the variables analysed in the review. The overwhelming majority of studies used a standardised measure to evaluate the presence of borderline personality disorder symptoms. However, a small number of studies relied on structured clinical interview by a psychiatrist, which is a less reliable method of diagnosis. In addition, three studies (11%) included participants who met some but not all diagnostic criteria (according to DSM-IV). While that improves the external reliability of individual studies, there may have been a range of severity of symptoms between studies. The decision to include studies using clinician diagnosis was specified in the protocol. Earlier reviews had not identified a large number of papers of brief interventions and at this stage we decided to keep exclusion criteria to a minimum to maximise the scope of the review. However, both of these selection factors may have increased heterogeneity in the participant population and could have contributed to the high levels of heterogeneity in some of the pooled data.

The purpose of this review was to evaluate how the research on brief psychological therapies for borderline personality disorder has developed, to provide more detailed guidance where possible on the use of such interventions for services, and to increase access to evidence based psychological interventions for service users. It was beyond the scope of the review to compare the effectiveness of brief interventions and those of longer-term therapies.

Conclusions

Interventions delivered in addition to ongoing support are more effective than those without. Mixed or group-only interventions can be considered, and group interventions may be particularly effective in reducing self-harm. However, brief psychological interventions overall have not been shown to be superior to planned, regular or structured contact. It is not possible to conclude whether interventions have an effect on self-harm or service use, or whether positive changes at the end of treatment remain stable over time. Despite the limitations of the pooled data, the review challenges

the guidance about avoiding the use of brief psychological interventions and highlights some important areas for future research.

References

*Andreasson, K., Krogh, J., & Wenneberg, C. (2016). Effectiveness of dialectical behaviour therapy versus collaborative assessment and management of suicidality treatment for reduction of self-harm in adults with borderline personality traits and disorder – A randomised observer-blinded clinical trial. Depression and Anxiety.33, 520-530.

*Andreoli, A., Burnand, Y., & Cochennec, M. (2016). Disappointed love and suicide: A randomised controlled trial of 'abandonment psychotherapy' among borderline patients. Journal of Personality Disorder. 30(2), 271-287.

American Psychiatric Association (2001). Practice guideline for the treatment of patients with borderline personality disorder. American Journal of Psychiatry. 158,1-52.

American Psychological Association (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, D.C: Author

Bateman, A. & Fonagy, P. (2009). Randomised controlled trial of outpatient mentalisation-based treatment versus structured clinical management for borderline personality disorder. American Journal of Psychiatry.1666, 1355-1364.

*Blum, N., St. John, D., & Pfohl, B. (2008). Systems training for emotional predictability and problem solving (STEPPS) for outpatients with borderline personality disorder: A randomised controlled trial and one-year follow-up. American Journal of Psychiatry.165(4), 468-478.

Bos, E., van Wel, E., Appelo, M. & Verbraak, M. (2010). A randomised controlled trial of a Dutch version of systems training for emotional predictability and problem solving for borderline personality disorder. Journal of Nervous and Mental Disease. 198(4), 299-304.

*Bos, E., van Wel, E., & Appelo, M. (2011). Effectiveness of systems training for emotional predictability and problem solving (STEPPS) for borderline personality disorder problems in a 'real world' sample: Moderation by diagnosis or severity? Psychotherapy and Psychosomatics. 80, 173-181

*Carter, G., Willcox, C., & Lewin, T. (2010). Hunter DBT project: A randomised controlled trial of dialectical behaviour therapy in women with borderline personality disorder. Australian and New Zealand Journal of Psychiatry.44, 162-173.

*Clarke, S., Thomas, P., & James, K. (2013). Cognitive analytic therapy for borderline personality disorder: A randomised controlled trial. British Journal of Psychiatry.202, 129-134.

Clarkin, J., Levy, K., Lenzenweger, M. & Kernberg, O. (2007). Evaluating three treatments for borderline personality disorder. American Journal of Psychiatry.164, 922-928.

Coid J, Yang M, & Tyrer P. (2006). Prevalence and correlates of personality disorder in Great Britain. British Journal of Psychiatry. 188, 423-431.

Crawford M.J., Price K, Gordon F, Josson M, Taylor B, Bateman A, Fonagy P, Tyrer P, & Moran P. (2009). Engagement and retention in specialist services for people with personality disorder. Acta Psychiatrica Scandinavia.119, 304–311.

*Crawford, M.J., Thana, L., Parker, J., Turner, O., Carney, A., McMurran, M., Moran, P., Weaver, T., Barrett, B., Roberts, S., Claringbold, A., Bassett, P., Sanatinia, R., Spong, A. (2020). Structured psychological support for people with personality disorder: feasibility randomised controlled

trial of a low intensity intervention. British Journal of Psychiatry Open. 6, e25, 1-9. Doi: 10.1192/bjo.2020.7.

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. JAMA Psychiatry.74(4), 319-328.

*Evans, K., Tyrer, P., & Catalan, J. (1999). Manual-assisted cognitive behaviour therapy (MACT): A randomised controlled trial of a brief intervention with bibliotherapy in the treatment of recurrent deliberate self-harm. Psychological Medicine.29, 19-2.

*Gratz, K. & Gunderson, J. (2006). Preliminary data on an acceptance-based emotion regulation group intervention for deliberate self-harm among women with borderline personality disorder. Behavior Therapy.37, 25-35.

*Gratz, K., Tull, M. Levy, R. (2014). Randomised controlled trial and uncontrolled 9-month follow-up of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. Psychological Medicine.44, 2099-2112.

Higgins, J.P., Altman, D.G., Gotzsche, P.C., Juni, P., Moher, D., Oxman A.D., Savovic, J., Schulz, K.F., Weeks, L., & Sterne, J.A.C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. British Medical Journal. 343. d5928.

*Huband, N., McMurran, M., & Evans, C. (2007). Social problem-solving plus psychoeducation for adults with personality disorder: Pragmatic randomised controlled trial. British Journal of Psychiatry.190, 307-313.

*Haeyen, S., van Hooren, S., van der Veld, W., Hutschemaekers, G. (2018). Efficacy of art therapy in individuals with personality disorder cluster B/C: A randomized controlled trial. Journal of Personality Disorders.32(4), 527-542.

*Jahangard, L., Haghighi, M., Bajoghli, H. (2012). Training emotional intelligence improves both emotional intelligence and depressive symptoms in inpatients with borderline personality disorder and depression. International Journal of Psychiatry in Clinical Practice.16, 197-204.

*Koons, C., Robins, C., & Tweed, J. (2001). Efficacy of dialectical behaviour therapy in women veterans with borderline personality disorder. Behavior Therapy.32, 371-390.

*Korrelboom, K., Marrisen, M., van Assendelft, T. (2011). Competative memory training (COMET) for low self-esteem in patient with personality disorders: A randomised effectiveness study. Behavioural and Cognitive Psychotherapy.39, 1-19.

*Kramer, U., Berger, T., Kolly, S. (2011). Effectives of motive-oriented therapeutic relationship in early-phase treatment of borderline personality disorder: A pilot study of a randomized trial. Journal of Nervous and Mental Disease.199(4), 244-250.

*Kramer, U., Kolly, S., & Berthoud, L. (2014). Effects of motive-oriented therapeutic relationship in a ten-session general psychiatric treatment of borderline personality disorder: A randomized controlled trial. Psychotherapy and Psychosomatics. 83, 176-186.

*Kramer, U. Pascual-Leone, A., Berthoud, L. (2016). Assertive anger mediates effects of dialectical behaviour-informed skills training for borderline personality disorder: A randomized controlled trial. Clinical Psychology and Psychotherapy.23, 189-202.

*Lin, T.-J., Huei-Chen, K., & Yung-Wei, W. (2018). The effectiveness of dialectical behaviour therapy skills training group vs. cognitive therapy group on reducing depression and suicide attempts for borderline personality disorder in Taiwan. Archives of Suicide Research.0, 1-18.

McMain, S., Links, P., Gnam, W., Guimond, T., Cardish, R., Korman, L., & Streiner, D. (2009) A randomised controlled trial of dialectical behaviour therapy versus general psychiatric management for borderline personality disorder. American Journal of Psychiatry.166, 1365-1374.

*McMain, S., Guimond, T., & Barnhart, R. (2017). A randomized trial of brief dialectical behaviour therapy skills training in suicidal patients suffering from borderline disorder. Acta Psychiatrica Scandinavia.135, 138-148.

McMurran, M., Huband, N., & Overton, E. (2010). Non-completion of personality disorder treatments: A systematic review of correlates, consequences, and interventions. Clinical Psychology Review. 30, 277–287. doi:10.1016/j.cpr.2009.12.002

*McMurran, M., Crawford, M., & Reilly, J. (2016). Psychoeducation with problem-solving (PEPS) therapy for adults with personality disorder: A pragmatic randomised controlled trial to determine the clinical effectiveness and cost-effectiveness of a manualised intervention to improve social functioning. Health Technology Assessment.20, Vol.52.

*Morton, J., Snowdon, S., Gopold, M. (2012). Acceptance and commitment group therapy treatment for symptoms of borderline personality disorder: A public sector pilot study. Cognitive and Behavioural Practice.19, 527-544.

National Health and Medical Research Council (2012). Clinical Practice Guideline for the Management of Borderline Personality Disorder. Melbourne: National Health and Medical Research Council.

National Institute for Health and Care Excellence (2009). Borderline personality disorder: Recognition and management. Clinical Guideline CG78.

Omar, H., Tejerina-Arreal, M., & Crawford, M. (2014). Are recommendations for psychological treatment of borderline personality disorder in current UK guidelines justified? Systematic review and subgroup analysis. Personality and Mental Health.8, 228-237.

Paris J. (2013). Stepped care: an alternative to routine extended treatment for patients with borderline personality disorder. Psychiatric Services.64, 1035-7.

Review Manager (RevMan) [Computer program] (2014). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration.

Simonsen S, Bateman A, Bohus M, Dalewijk HJ, Doering S, Kaera PM, Moran P, Renneberg B, Soler Ribaudi J, Taubner S, Wliberg T, & Mehlum L. (2019). European guidelines for personality disorders: past, present, and future. Bordering Personality Disorder and Emotion Dysregulation.6(9), 1-10.

*Soler, J., Pascual, J., & Tiana, T. (2009). Dialectical behaviour therapy skills training compared to standard group therapy in borderline personality disorder: A 3-month randomised controlled clinical trial. Behaviour Research and Therapy 47, 353-358.

*Springer, T., Lohr, N., Buchtel, H. (1996). A preliminary report of short-term cognitive-behavioural group therapy for inpatients with personality disorders. Journal of Psychotherapy and Practice Research.5, 57-71.

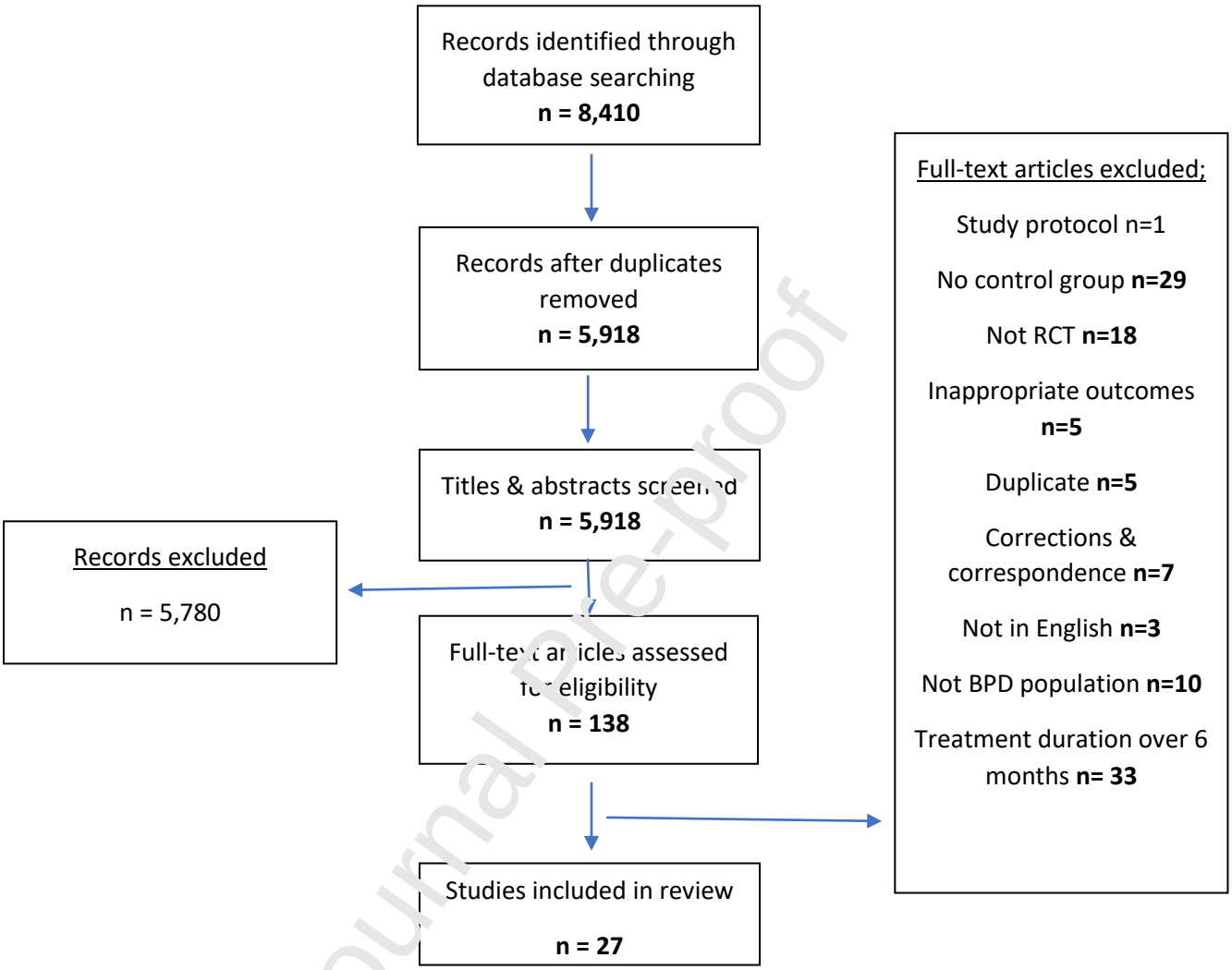
Stoffers-Winterling JM, Völlm BA, Rücker G. Timmer, A., & Lieb, K. (2012). Psychological therapies for people with borderline personality disorder. Cochrane Database of Systematic Reviews. Issue 8. Art. No.: CD005652. DOI: 10.1002/14651858.CD005652.pub2

*Weinberg, I., Gunderson, J., & Hennen, J. (2006). Manual assisted cognitive treatment for deliberate self-harm in borderline personality disorder patients. Journal of Personality Disorder.20(5), 482-492.

*Zanarini, M. & Frankenburg, F. A. (2008). preliminary, randomized trial of psychoeducation for women with borderline personality disorder. Journal of Personality Disorders.22(3), 284-290.

*Zanarini, M., Conkey, L., Temes, C. (2018). Randomized controlled trial of web-based psychoeducation for women with borderline personality disorder. Journal of Clinical Psychiatry.79(3), 52-59.

Figure 1. PRISMA flow diagram



Author biography

Amanda J. Spong

I qualified as a clinical psychologist in 2000 from the University of East Anglia doctorate course. I have worked in NHS mental health services since this time to the present date, in community and inpatient services. In 2017 I was granted a NIHR CLAHRC research fellowship whilst working as the Principal Clinical Psychologist and personality disorder pathway lead for the community personality disorder service in Cambridgeshire & Peterborough NHS trust. This systematic review was carried out with the support of this fellowship and focused on a clinically and service relevant area; namely to improve the evidence base needed to increase access to psychological interventions for patients with borderline personality disorder. The research was supervised by Professor Peter B. Jones because of his expertise in systematic reviews and meta-analyses.

Author Disclosure Statements

Role of Funding Sources

This paper is a summary of independent research part-funded by the National Institute of Health Research (NIHR) Applied Research Collaboration East of England (ARC EoE) at Cambridgeshire & Peterborough NHS Foundation Trust. We are grateful for their support. The views expressed are those of the authors and not necessarily those of the NIHR, the NHS or the Department of Health and Social Care.

Contributors

Author A.S conceptualised the study and I.C.H.C., J.G., P.B.J. and M.J.C. advised on the development of the study protocol. A.S. and I.C.H.C. selected the searches, extracted the results, and conducted risk of bias analysis. A.S. conducted the meta-analysis and analysis of publication bias, with support and advice from J.G., and wrote the first draft of the manuscript. All authors have seen and approved the final version.

Conflict of Interest

All authors declare they have no conflicts of interest.

Acknowledgements

We would like to thank Ms Isla Kuhn, Head of Medical Library Services at the University of Cambridge Clinical School, for her contribution to the design and running of the database searches.

Highlights

- First review to evaluate brief psychological interventions for borderline personality disorder
- Additional support is needed alongside interventions to maximise change
- Planned, generic support may be as effective as specialist interventions for borderline symptoms and social functioning
- Analysis of the impact of mode of delivery of interventions was inconclusive
- Follow-up data is lacking across all outcomes; it is not clear if symptomatic change is sustained