Published in final edited form as:

Acta Psychiatr Scand. 2009 July ; 120(1): 53–61. doi:10.1111/j.1600-0447.2008.01340.x.

Predictors of response to Systems Training for Emotional Predictability and Problem Solving (STEPPS) for borderline personality disorder: an exploratory study

D. W. Black, J. Allen, D. St. John, B. Pfohl, B. McCormick, and N. Blum

Department of Psychiatry, University of Iowa Roy J. and Lucille A. Carver College of Medicine, Iowa City, IA, USA

Abstract

Objective—Few predictors of treatment outcome or early discontinuation have been identified in persons with borderline personality disorder (BPD).

Aim—The aim of the study was to examine the relationship between baseline clinical variables and treatment response and early discontinuation in a randomized controlled trial of System Training for Emotional Predictability and Problem Solving, a new cognitive group treatment.

Method—Improvement was rated using the Zanarini Rating Scale for BPD, the Clinical Global Impression Scale, the Global Assessment Scale and the Beck Depression Inventory. Subjects were assessed during the 20 week trial and a 1-year follow-up.

Results—Higher baseline severity was associated with greater improvement in global functioning and BPD-related symptoms. Higher impulsivity was predictive of early discontinuation. Optimal improvement was associated with attending 15 sessions.

Conclusion—Subjects likely to improve have the more severe BPD symptoms at baseline, while high levels of impulsivity are associated with early discontinuation

Keywords

group therapy; cognitive-behavioral treatment; Systems Training For Emotional Predictability and Problem Solving; borderline personality disorder; response predictors

Introduction

Borderline personality disorder (BPD) impairs quality of life, is associated with excessive health care utilization, and leads to suicide in 4–10% of patients (1–3). The disorder affects nearly 2% of the general population, and as many as 15% of psychiatric out-patients (4–6). Treatment is challenging and results are generally modest with medication (7, 8). Psychotherapy research has been actively pursued and several treatment models are now evidence-based including dialectical behavior therapy (9, 10), cognitive-behavioral therapy (11), schema-focused therapy (12), transference-focused psychotherapy (13) and the mentalization therapy of Bateman and Fonagy (14, 15). These programs have recently been joined by Systems Training for Emotional Predictability and Problem Solving (STEPPS), a

Copyright © 2009 The Authors

Dr Black, 2-126b MEB /Psychiatry Research, University of Iowa Carver College of Medicine, Iowa City, IA 52242, USA. donald-black@uiowa.edu.

Declaration of interest: None.

group treatment developed at Iowa and recently found effective across a spectrum of illness-specific and global outcome measures in a randomized controlled trial (RCT) of borderline out-patients (16). STEPPS combines cognitive-behavioral elements with skills training and a systems component for family members, friends and those with whom the patient regularly interacts (17, 18).

Despite the development of these programs, little progress has been made in identifying predictors of treatment response to either psychotropic medication or psychotherapy in persons with BPD. Similarly, little effort has been made in identifying predictors of treatment drop-out. The latter problem is particularly germane because patients with BPD are at risk for prematurely discontinuing treatment.

The literature provides few clues to predicting patient outcome, in part because of the small samples involved that precluded analysis. In reviewing follow-up study results, Lieb et al. (3) report that poor outcome is associated with worse baseline psychopathology, affective instability, increased length of prior hospitalizations, the presence of dysphoria, a family history of mental illness, younger age at entering treatment, maternal psychopathology, history of parental brutality and childhood sexual abuse. Nonetheless, these variables have been of limited value in aiding clinicians because they are non-specific and general.

Also relevant are the reports of Soloff et al. (19) and Verheul et al. (20). In the former, Soloff et al. (19) found that severity of schizotypal symptoms, hostility and suspiciousness predicted favorable response to haloperidol, yet schizotypal symptoms and paranoia predicted poor response to amitriptyline; severe 'character' traits predicted poor response in the placebo-treated group. In the later, Verheul et al. (20) reported that dialectical behavior therapy was superior to treatment as usual among patients with 'high-severity' determined by a median split on lifetime number of self-mutilating acts.

Almost no work has been done to identify predictors of early study discontinuation. In a relevant report of psychodynamic psychotherapy, Smith et al. (21) found that younger persons and those with higher levels of hostility were more likely to discontinue therapy prematurely. In a study of a social problem-solving treatment for personality disordered individuals, Huband et al. (22) reported that subjects in a 'high attrition' group were more likely to have a forensic history, to have personality disorders in more than one cluster, and to have greater baseline impulsivity scores. Distance from the facility and treatment site were not related to attrition.

For this analysis, we assessed predictors of treatment response and premature study discontinuation in persons with BPD who participated in our RCT of the STEPPS program (16).

Aims of the study

We hypothesized that variables associated with improvement would include persons with higher symptom severity, a higher number of abnormal personality traits and a lower level of impulsivity. We did not expect age, gender, history of suicidal and self-harm behaviors, or past substance abuse to predict response. We also hypothesized that a higher number of personality disorders and greater impulsivity also would predict early discontinuation.

Material and methods

Subjects

Subjects 18 years or older who met DSM-IV criteria (23) for BPD were recruited for the RCT. The diagnosis was confirmed using the Structured Interview for DSM-IV Personality

(SIDP-IV) (24). Subjects gave written informed consent according to procedures approved by the University of Iowa Institutional Review Board. They could not have a diagnosis of schizophrenia, schizoaffective disorder, psychotic mood disorder or a primary neurological disorder; have obvious cognitive impairment; have current (past month) substance abuse or dependence (except tobacco dependence); or have participated in a STEPPS program.

Assessments

Axis I and II comorbidity was assessed with the Structured Clinical Interview for DSM-IV (25) and the SIDP-IV. Baseline ratings included the Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD) (26), to assess anger, mood reactivity, emptiness, identity disturbance, stress-related paranoid ideation / dissociation, efforts to avoid abandonment, suicidal / self-harm behavior, impulsivity and unstable relationships; the Borderline Evaluation of Severity Over Time (BEST) (27), a self-report scale to assess symptoms of BPD; the Symptom Checklist-90-R (SCL-90-R) (28), a self-report instrument to assess a wide range of psychiatric symptoms; the Social Adjustment Scale (SAS) (29), a self-report instrument to rate subject's social functioning in five domains including work, social and leisure activities, relationship with extended family, marital and parental role, and economic dependence. The Clinical Global Impression (CGI) severity scale (30), the Barratt Impulsiveness Scale (BIS) (31), the Global Assessment Scale (GAS) (32) and the Beck Depression Inventory (BDI) (33) were also used in the analysis. We recorded the types and amounts of professional and mental health and medical treatment a person had received, as well as past suicide attempts and episodes of deliberate self-harm. We calculated the distance subjects lived from the University of Iowa Hospitals and Clinics where the therapy sessions were held.

Description of treatment as usual and the STEPPS program

Subjects were randomly assigned to STEPPS plus treatment as usual or to treatment as usual alone. All subjects were encouraged to continue their usual care including individual psychotherapy, medication and case management. Subjects assigned to treatment as usual alone were not allowed to attend a STEPPS program until they completed the 20 week study. Subjects were followed for 1 year after completing the 20 week study period.

STEPPS is a 20-week manual-based group treatment program for out-patients with BPD (34). The program combines cognitive-behavioral elements with skills training. There are three main components: i) psychoeducation about BPD, ii) emotion management skills training and iii) behavior management skills training. The first component teaches subjects to replace misconceptions about BPD with an awareness of the thoughts, feelings and behaviors that define it and to identify their own schemas (i.e. cognitive filters) that drive their behaviors. The second component teaches skills to better manage the cognitive and emotional effects of BPD: distancing, communicating, challenging, distracting and managing problems. The third component teaches behavioral skills that subjects are encouraged to master: goal setting, healthy eating behaviors, sleep hygiene, regular exercise, leisure activities, health monitoring (e.g. medication adherence), avoiding self-harm and interpersonal effectiveness.

For this analysis, predictors of treatment response were examined for both treatment groups. Because treatment as usual was allowed for both groups, the predictors suggest patient characteristics associated with improvement in BPD-related symptoms for subjects receiving usual care (individual psychotherapy, medication and case management). As described in the next section, we also attempted to identify factors that predicted response differentially for the two treatment groups; these predictors suggest patient characteristics whose effect on improvement is modified by the STEPPS program.

Statistical analysis

Subjects were assessed at baseline, at regular intervals during the treatment period (weeks 4, 8, 12, 16 and 20), and at regular intervals during 1 year of follow-up (1, 3, 6, 9 and 12 months). We assessed treatment response with five rating scales: the CGI, ZAN-BPD, BEST, GAS and BDI. Using these scales, we measured improvement in three domains: overall BPD severity (ZAN-BPD, BEST), global severity (CGI, GAS) and depressive symptoms (BDI). For the CGI, we classified as responders those subjects who were rated as 'much improved' or 'very much improved' at any assessment beyond the first 10 weeks of the treatment program. For the other rating scales, we defined level of treatment response as the improvement relative to the baseline assessment. For each subject, scale scores were regressed on time (weeks since baseline assessment) and the resulting slope estimates were used as the measures of improvement.

Number of STEPPS sessions attended and last follow-up time are both indicators of early discontinuation. For our analyses, we considered last assessment time as the measure of early discontinuation because it was applicable to both treatment groups. Among those assigned to STEPPS plus treatment as usual, 51% attended at least 10 sessions and 39% attended at least 15. Thirtythree per cent of those assigned to STEPPS plus treatment as usual and 22% assigned to treatment as usual were assessed through the 1-year follow-up, thus had last follow-up time of week 72. The median last follow-up time was week 33 for those assigned to STEPPS plus treatment as usual and week 20 for those assigned to treatment as usual.

Baseline predictors included measures of BPD symptom severity (ZAN-BPD total and its four subscales assessing affective disturbance, cognitive disturbance, impulsivity and disturbed relationships, BEST), overall symptom severity (CGI severity, GAS, SCL-90-R), depression (BDI), impulsivity (BIS), social adjustment (SAS), important demographic variables (gender, age, education level and distance traveled from home to treatment center), treatment received (out-patient mental health visits in last month, individual psychotherapy, past psychiatric hospitalization, number of current psychotropic medications), prior suicide and self-harm attempts, axis I psychiatric comorbidity (current depression, current anxiety disorder, lifetime substance abuse and lifetime eating disorder) and personality disorders (antisocial, avoidant, dependent, narcissistic, obsessive-compulsive, paranoid and total number of axis II disorders). Number of STEPPS sessions attended was also used as a predictor for the STEPPS plus treatment as usual group. Because greater improvement has already been observed for the STEPPS plus treatment as usual group, treatment group was also used as a predictor variable.

Our strategy was to first determine which baseline predictors were individually predictive of each response measure and then to determine which baseline predictors were jointly predictive of each response measure in multiple regression models. Because the factors that affect treatment response may depend on aspects of the treatment received, predictor-by-treatment group interactions were tested in the multiple regression models. Non-significant interactions were dropped from the models. For all analyses, we used a significance level of 0.05. Statistical analyses were carried out using sas, version 8 (35).

For the dichotomous outcome (CGI response), we compared the means (responders vs. non-responders) of the continuous baseline predictors using the Student's *t*-test; we compared the proportions of the dichotomous baseline predictors using Pearson's chi-squared test. Effect sizes (Cohen's *d* for the continuous predictors, the phi coefficient for the dichotomous predictors) were also reported. The baseline predictors that were significantly different by response status were then entered together as predictors of CGI response in a multiple

logistic regression model. Interactions of treatment group with each predictor were tested. The multiple logistic regression analysis was carried out with the LOGISTIC procedure.

For the continuous measures of improvement (ZAN-BPD, BEST, BDI and GAS) and last assessment time, we correlated each measure with each baseline predictor. For each measure, the baseline predictors that had significant correlations entered together in a multiple linear regression model. Again, predictor-by-treatment group interactions were tested. The multiple linear regression analyses were carried out with the REG procedure. In a secondary analysis, we sought to determine a 'threshold' for the effect of attending STEPPS. We compared mean improvement for subjects who had attended at least 10 STEPPS sessions to the mean improvement for subjects who had attended fewer than 10 STEPPS sessions. This analysis was then repeated using 15 sessions as the cutoff. This analysis was conducted for two of the response measures: improvement in ZAN-BPD total and GAS.

Results

The results are based on an analysis of 164 subjects who were randomly assigned to STEPPS plus treatment as usual ($n = 92$) or to treatment as usual alone ($n = 72$). Most subjects (85%) were women. One subject assigned to STEPPS plus treatment as usual was dropped from the analysis because of the presence of a chronic neurological condition not evident at intake.

Individual predictors of improvement

In Table 1, we present tests for differences in baseline predictor variables between treatment responders (subjects who received a CGI rating of 'much improved' or 'very much improved' at week 12 or a later assessment) and non-responders. For continuous predictors, the means are given for responders and non-responders and the t -test is used to test for significant differences. Cohen's d statistic represents the difference in the two means, divided by the pooled standard deviation. Because the continuous baseline predictors are not coded in a consistent fashion, we adjusted the direction of the d statistic so that the meaning was consistent across predictors. For example, the mean baseline ZAN-BPD affective disturbance score was higher for responders but because higher scores imply greater severity, the d statistic is negative. Thus, positive d values suggest that the responders had less severe symptoms at baseline. For the dichotomous predictors, the proportions are given for responders and non-responders and the chi-squared test is used to test for significant differences in the proportions. The phi coefficient is presented as an effect size measure. From Table 1, we see that responders had significantly higher ZAN-BPD affective disturbance scores at baseline, were more likely to be randomized to STEPPS plus treatment as usual, and were more likely to be men. Within the STEPPS plus treatment as usual group, responders had attended significantly more STEPPS sessions.

In Table 2, we present the statistical tests of associations between baseline predictor variables and the four improvement measures (slopes for ZAN-BPD total score, BEST, BDI and GAS). We present correlation coefficients, where the direction of the coefficient is adjusted so that the coefficients have a consistent meaning for the predictors measuring baseline severity: positive correlations suggest that lesser severity is associated with greater improvement. Improvement in ZAN-BPD total score was associated with greater global severity (CGI Severity), greater severity of BPD-related symptoms (all ZAN-BPD scales), more personality disorders, being randomized to STEPPS plus treatment as usual, having attempted to harm oneself in the past and having paranoid personality disorder. Improvement in BEST total score was associated with greater severity of BPD-related symptoms (BEST total), being randomized to STEPPS plus treatment as usual and having a

lifetime substance use disorder. Improvement in GAS score was associated with greater global severity (i.e. worse GAS score) at intake and being randomized to STEPPS plus treatment as usual.

Multiple regression models for improvement

In Table 3, we present the results of the multiple logistic regression model for response (as determined by whether a subject was rated as 'much improved' or 'very much improved' at the week 12 assessment or later). Earlier, we saw that the responders had higher ZAN-BPD affective disturbance scores at baseline, were more likely to be men and were more likely to be randomized to STEPPS plus treatment as usual. These predictors entered the multiple logistic regression model, along with treatment group-by-predictor interactions. The interaction terms were not significant and were removed from the model. The remaining three predictors remained significant in the multiple logistic regression model. Thus, treatment response was jointly associated with treatment group, male gender and having greater initial severity of BPD-related symptoms, as measured by the ZAN-BPD affective disturbance subscale score.

In Table 4, we present the results of the multiple linear regression models for the improvement measures. The predictors that had significant correlations (Table 2) entered the multiple linear regression model, along with treatment group-by-predictor interactions. The baseline ZAN-BPD total score and each of its subscales were negatively correlated with improvement in ZAN-BPD total scale. Because these predictors are highly correlated with one another, we only entered the baseline ZAN-BPD total score into the multiple linear regression model. Similarly, both paranoid personality disorder and total number of personality disorders were positively correlated with improvement in ZAN-BPD total scale. Because these two predictors are highly correlated, we only entered total number of personality disorders into the multiple linear regression model. In the model for improvement in ZAN-BPD total score, the interactions of number of personality disorders and baseline ZAN-BPD total score with treatment group were significant. The results suggest that number of personality disorders is positively associated with ZAN-BPD improvement for the STEPPS plus treatment as usual group, but not for the treatment as usual alone group. Greater baseline severity of BPD-related symptoms (baseline ZAN-BPD total score) was associated with ZAN-BPD improvement for the treatment as usual alone group, and to a lesser extent for the STEPPS plus treatment as usual group. Past self-harm acts and baseline CGI severity were not predictive of ZAN-BPD improvement in the multiple predictor model.

Greater baseline severity of BPD-related symptoms (BEST total score) and being randomized to STEPPS plus treatment as usual were significant predictors in the model for improvement in BEST total score. Having a lifetime substance use disorder was not predictive in the regression model, nor was the predictor-by-treatment group interaction terms. Treatment group was the only predictor variable significantly correlated with improvement in BDI score; thus, a multiple linear regression model was not warranted for this measure of improvement. Greater baseline severity of global symptoms (GAS score) and treatment group were predictive of improvement in GAS score. The baseline GAS score by treatment group interaction was not significant. These results suggest that the predictors of improvement do not vary much by treatment group.

Predictors of discontinuation

Greater levels of impulsivity at baseline assessed with the BIS and fewer psychotropic medications taken were positively correlated with early discontinuation (last assessment time). Naturally, number of STEPPS sessions attended was also positively correlated with

last assessment time within the STEPPS plus treatment as usual group (Table 2). In the multiple predictor model (Table 4), greater impulsivity remained predictive of early discontinuation, whereas number of psychotropic medications was not. The predictor-by-treatment group interactions were not significant.

Threshold for the effect of attending more STEPPS sessions

The results suggest that 15 STEPPS sessions may be a threshold for achieving optimal improvement. For STEPPS plus treatment as usual subjects who attended at least 15 sessions, improvement in ZAN-BPD was greater than for subjects who attended fewer than 15 sessions (Cohen's $d = 0.24$), although the result was not statistically significant ($P = 0.375$). For subjects who attended 10 sessions, improvement in ZAN-BPD was only slightly greater (Cohen's $d = 0.08, P = 0.783$) than for subjects who attended fewer than 10 sessions. The same trend held for improvement in GAS score. For subjects who attended at least 15 sessions, improvement was greater (Cohen's $d = 0.29, P = 0.264$); for subjects who attended 10 sessions, improvement was not affected (Cohen's $d = -0.04, P = 0.268$).

Discussion

We approached the search for response predictors with the understanding that both treatment groups might have unique predictors, and that the predictors might vary depending upon the measure of response used, yet predictors largely overlapped. The most consistent finding was that greater initial clinical severity predicted better response. This was true for both treatment groups and for all response domains, with the exception of improvement in depressive symptoms measured with the BDI. This finding suggests that BPD patients with greater initial severity stand to gain the most from treatment and are more likely than others to experience improvements in global severity and BPD-related symptoms. This finding seems counterintuitive, yet could be explained by the phenomenon of 'regression to the mean' in which subjects at the end of a frequency distribution have the farthest to move. Another possibility is that the results could be due to state personality effects in which a personality disordered subject appears more ill under stress and is rated as being more ill that may be true.

Contrary to expectation, older age was not associated with improved response for either group. Men were more likely to receive CGI ratings of 'much improved' or 'very much improved', but were no more likely to show improvement on the other response measures. Prior acts of self-harm and having a lifetime substance use disorder were associated with improvement in BPD-related symptoms (ZAN-BPD and BEST total score respectively), but neither was predictive in the multiple regression models. The associations may have been a byproduct of the relationship of prior self-harm acts, prior substance misuse and initial severity.

The multiple regression analyses suggest that few of the predictors of response vary by type of treatment received (treatment as usual alone or STEPPS plus treatment as usual). Thus, while greater improvement was observed for patients randomized to STEPPS plus treatment as usual, there was little evidence of interactions between STEPPS and the patient characteristics associated with improvement in BPD-related symptoms.

There were few predictors of early discontinuation, a problem that vexes clinicians and researchers alike. Greater impulsivity and taking fewer psychotropic medications were associated with early discontinuation, but only greater impulsivity remained predictive in the multiple regression model. Unlike Smith et al. (21) who studied drop-outs from psychodynamic therapy, we did not find a significant relationship with age ($r = 0.10, P = 0.20$) and were unable to assess the effect of hostility to dropping out as we had no direct

measure of this variable. Surprisingly, distance traveled from home to the treatment center was not correlated with early discontinuation ($r = -0.01$). We had assumed that the farther our patients had to drive to attend the program, the greater their likelihood of dropping out, but this turned out not to be true.

Optimal number of sessions

The results of our secondary analysis suggest that attending 15 or more STEPPS sessions produces optimal benefit, although subjects who attended fewer sessions also benefited. Of the 65 subjects assigned to STEPPS plus treatment as usual, 35 (54%) attended 15 or more sessions. Hence, many did not reach the threshold value. The efficacy of the STEPPS program might have been greater had more subjects completed the program.

Limitations

There are several limitations to the analysis. First, data were collected in the course of a clinical trial and several potentially important variables were not assessed (e.g. childhood sexual abuse, hostility). Second, we experienced a relatively high discontinuation rate that could have affected the results, in that those who dropped out might have been less likely to improve. The small number of men and minorities did not allow us to examine whether response predictors vary by subgroup. Another limitation was that many statistical tests were performed, resulting in greater opportunity for type 1 errors. For that reason, this study should be viewed as exploratory. Finally, the fact that we did not seek subjects with recent suicidal or self-harm acts limits comparison with studies that actively recruited such persons (9, 10). Future studies of predictors of treatment response in BPD patients are needed to extend and replicate our findings. This will allow the field to better understand patient characteristics that may be associated with improvement.

Acknowledgments

This study was supported by the National Institute of Mental Health, Bethesda, MD (MH63746). The authors are grateful to Joann Franklin and Becky Hansel for data collection and data entry.

References

1. Bender DS, Dolan RT, Skodol AE, Sanislow CA, Dyck IR, McGlashan TH. Treatment utilization by patients with personality disorders. *Am J Psychiatry*. 2001; 158:295–302. [PubMed: 11156814]
2. Black DW, Blum N, Pfohl B, Hale N. Suicidal behavior in borderline personality disorder: prevalence, risk factors, prediction, and prevention. *J Personal Disord*. 2004; 18:226–239.
3. Lieb K, Zanarini MC, Schmahl C, Linehan MM, Bohus M. Borderline personality disorder. *Lancet*. 2004; 364:453–461. [PubMed: 15288745]
4. Swartz M, Blazer D, George L, Winfield I. Estimating the prevalence of borderline personality disorder in the community. *J Personal Disord*. 1990; 4:257–272.
5. Torgerson S, Kringlen E, Cramer V. The prevalence of personality disorders in a community sample. *Arch Gen Psychiatry*. 2001; 58:590–596. [PubMed: 11386989]
6. Widiger, TA.; Frances, AJ. Epidemiology, diagnosis, and comorbidity of borderline personality disorder. In: Tasman, A.; Hales, RE.; Frances, AJ., editors. *American psychiatric press review of psychiatry*. Washington, DC: American Psychiatric Press; 1989. p. 8-24.
7. Soloff PH. Psychopharmacology of borderline personality disorder. *Psychiatr Clin North Am*. 2000; 23:169–191. [PubMed: 10729338]
8. Zanarini MC. Update on pharmacotherapy of borderline personality disorder. *Curr Psychiatry Rep*. 2004; 6:66–70. [PubMed: 14738708]
9. Linehan MM, Armstrong HE, Suarez A, Allmon D, Heard HL. Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Arch Gen Psychiatry*. 1991; 48:1060–1064. [PubMed: 1845222]

10. Linehan MM, Comtois KA, Murray AM, et al. Two year randomized controlled trial and follow up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. *Arch Gen Psychiatry*. 2006; 63:757–766. [PubMed: 16818865]
11. Davidson K, Norrie J, Tyrer P, et al. The effectiveness of cognitive behavior therapy for borderline personality disorder: results from the borderline personality disorder study of cognitive therapy (BOSCOT) trial. *J Personal Disord*. 2006; 20:450–465.
12. Giesen-Bloo J, van Dyck R, Spinhoven P, et al. Outpatient psychotherapy for borderline personality disorder – randomized trial of schema-focused therapy vs transference- focused psychotherapy. *Arch Gen Psychiatry*. 2006; 63:649–658. [PubMed: 16754838]
13. Clarkin JF, Levy KN, Lenzenweger MF, Kernberg OF. Evaluating three treatments for borderline personality disorder: a multi-wave study. *Am J Psychiatry*. 2007; 164:922–928. [PubMed: 17541052]
14. Bateman A, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder – a randomized controlled trial. *Am J Psychiatry*. 1999; 156:1563–1569. [PubMed: 10518167]
15. Bateman A, Fonagy P. 8-year follow-up of patients treated for borderline personality disorder: mentalization-based treatment versus treatment as usual. *Am J Psychiatry*. 2008; 165:631–638. [PubMed: 18347003]
16. Blum N, Pfohl B, St John D, et al. Systems Training for Emotional Predictability and Problem Solving (STEPPS) for outpatients with borderline personality disorder: a randomized controlled trial and 1-year follow-up. *Am J Psychiatry*. 2008; 165:468–478. [PubMed: 18281407]
17. Black DW, Blum N, Pfohl B, St John D. The STEPPS group treatment program for outpatients with borderline personality disorder. *J Contemp Psychother*. 2004; 34:193–210.
18. Blum N, Pfohl B, St John D, Monahan P, Black DW. STEPPS: a cognitive behavioral systems-based group treatment for outpatients with borderline personality disorder – a preliminary report. *Compr Psychiatry*. 2002; 43:301–310. [PubMed: 12107867]
19. Soloff PH, George A, Nathan S, et al. Amitriptyline versus haloperidol in borderlines: final outcomes and predictors of response. *J Clin Psychopharmacol*. 1989; 9:238–246. [PubMed: 2768542]
20. Verheul R, Van den Bosch LMC, Koeter MWJ, De Ridder MAJ, Stijnen T, Van den Brink W. Dialectical behavior therapy for women with borderline personality disorder: 12-month randomized clinical trial in the Netherlands. *Br J Psychiatry*. 2003; 182:135–140. [PubMed: 12562741]
21. Smith TE, Koenigsberg HW, Yeomans FE, Clarkin JF, Selzer MA. Predictors of dropout in psychodynamic psychotherapy of borderline personality disorder. *J Psychother Pract Res*. 1995; 4:205–213. [PubMed: 22700251]
22. Huband N, McMurran M, Evans C, Duggan C. Social problem-solving plus psychoeducation for adults with personality disorder. *Br J Psychiatry*. 2007; 190:307–313. [PubMed: 17401036]
23. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: American Psychiatric Association; 1994.
24. Pfohl, B.; Blum, N.; Zimmerman, M. Structured interview for DSM-IV personality (SIDP-IV). Washington, DC: American Psychiatric Press; 1997.
25. Spitzer, RL.; Williams, JBW.; Gibbon, M. Structured Clinical Interview for DSM-IV. New York: New York State Psychiatric Institute, Biometrics Research; 1994.
26. Zanarini, MC.; Frankenburg, FR. Zanarini Rating Scale for Borderline Personality Disorder (ZAN-BPD). Belmont, MA: Harvard Medical School; 2001.
27. Pfohl B, Blum N, McCormick B, StJohn D, Allen J, Black DW. Reliability and validity of the Borderline Evaluation of Severity Over Time: a new scale to measure severity and change in borderline personality disorder. *J Person Disord*. (in press).
28. Derogatis, LR. SCL-90-R administration, scoring, and procedures manual II. Towson, MD: Clinical Psychometric Research; 1983.
29. Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry*. 1976; 33:1111–1115. [PubMed: 962494]

30. Guy, W. ECDEU assessment manual for psychopharmacology. Washington, DC: US Department of HEW Publications; 1976. p. 76-338.p. 217-222.
31. Barratt E. Anxiety and impulsiveness related to psychomotor efficiency. *Percept Mot Skills*. 1959; 9:191– 198.
32. Endicott J, Spitzer RL, Fleiss JL, Cohen J. The Global Assessment Scale: a procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry*. 1976; 33:766–771. [PubMed: 938196]
33. Beck, AT. Depression inventory. Philadelphia, PA: Philadelphia Center for Cognitive Therapy; 1978.
34. Blum, N.; Bartels, N.; StJohn, D.; Pfohl, B. Systems Training for Emotional Predictability and Problem Solving (STEPPS) – group treatment for borderline personality disorder. Coralville, IA: Blum's Books; 2002.
35. SAS. SAS / STAT user's guide: version 8. Cary, NC: SAS Institute Inc; 1999.

Significant outcomes

- Higher baseline severity was associated with greater improvement in functioning and clinical symptoms.
- Higher levels of impulsivity were predictive of early discontinuation.
- Optimal levels of improvement were associated with attending 15 or more System Training for Emotional Predictability and Problem Solving sessions.

Limitations

- Data were collected in the course of a randomized controlled trial and not all potentially important variables were assessed, such as hostility or childhood adversity.
- The discontinuation rate was high.
- Few men and minorities participated in the study limiting generalizations to these groups.

Table 1

Means of baseline predictors by response status

Baseline predictor variables	Mean		Student's <i>t</i> -test		
	R	NR	<i>d</i>	<i>t</i>	<i>P</i>
Continuous variables					
CGI severity [†]	5.05	4.95	-0.13	-0.70	124 0.487
GAS score	38.48	40.47	-0.18	0.94	122 0.350
ZAN-BPD total [†]	19.40	17.01	-0.33	-1.88	123 0.063
Affective [†]	8.28	6.79	-0.52	-2.94	123 0.004
Cognitive [†]	4.15	3.97	-0.08	-0.43	123 0.671
Disturbed [†]	4.11	3.71	-0.19	-1.07	123 0.285
Impulsivity [†]	2.89	2.59	-0.14	-0.78	123 0.437
BDI total [†]	30.06	30.75	0.06	0.31	119 0.761
BEST total [†]	40.51	40.32	-0.02	-0.10	117 0.920
BIS total [†]	78.41	79.40	0.08	0.40	115 0.687
SAS total [†]	2.81	2.81	0.01	0.03	118 0.976
SCL-90-R Global Severity Index [†]	16.14	17.00	0.12	0.66	118 0.508
Age	33.19	30.76	0.26	-1.41	124 0.162
Education level achieved	4.68	4.29	0.27	-1.39	124 0.168
Distance traveled to clinic	32.90	30.01	0.06	-0.41	124 0.680
Out-patient mental health visits	3.60	5.72	-0.29	1.44	115 0.154
No. psychotropic medications taken	3.32	3.56	-0.09	0.49	124 0.626
No. STEPPS sessions attended	14.13	9.44	0.67	-3.13	63 0.003
No. personality disorders	2.98	3.54	-0.31	1.70	124 0.091
Dichotomous variables					
	Proportion		Chi-squared test		
	R	NR	phi	χ^2	<i>P</i>
Treatment group (STEPPS plus TAU)	0.81	0.34	0.45	25.70	1 0.000
Gender (male)	0.28	0.11	0.21	5.41	1 0.020
Individual psychotherapy participation	0.59	0.60	-0.01	0.01	1 0.924

Baseline predictor variables		Mean		Student's <i>t</i> -test		
Continuous variables		R	NR	<i>d</i>	<i>t</i>	<i>P</i>
Past psychiatric history		0.71	0.75	-0.04	0.17	1
Past self-harm acts		0.66	0.73	-0.08	0.79	1
Past suicide attempt(s)		0.74	0.75	0.00	0.00	1
Current major depression		0.57	0.43	0.14	2.45	1
Current anxiety disorder		0.81	0.80	0.01	0.02	1
Lifetime anxiety disorder		0.83	0.86	-0.04	0.22	1
Lifetime substance use disorder		0.62	0.61	0.01	0.01	1
Lifetime eating disorder		0.32	0.43	-0.11	1.53	1
Antisocial PD		0.23	0.24	-0.01	0.01	1
Avoidant PD		0.45	0.57	-0.12	1.78	1
Dependent PD		0.21	0.35	-0.15	2.81	1
Narcissistic PD		0.15	0.22	-0.08	0.84	1
Obsessive-compulsive PD		0.55	0.52	0.03	0.14	1
Paranoid PD		0.30	0.34	-0.05	0.26	1

R, mean (proportion) for responders; NR, mean (proportion) for non-responders, responders defined as subjects who were rated 'much improved' or 'very much improved' at week 12 or later assessment; *d*, Cohen's *d* (difference in means in standard deviation units); phi, phi coefficient; PD, personality disorder; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder; BEST, Borderline Evaluation of Severity Over Time; BDI, Beck Depression Inventory; GAS, Global Assessment Scale; BIS, Barratt Impulsiveness Scale; SAS, Social Adjustment Scale; SCL-90-R, Symptom Checklist-90-R; STEPPS, Systems Training for Emotional Predictability and Problem Solving.

[†]Higher scores imply greater symptom severity (less favorable).

Table 2

Correlations of baseline predictors and improvement measures

Baseline predictor variables	Improvement measures				
	ZAN-BPD	BEST	BDI	GAS	LAT
Continuous variables					
CGI severity [†]	-0.30 [*]	-0.06	0.10	-0.17	-0.07
GAS score	-0.17	-0.03	0.09	-0.34 [*]	0.03
ZAN-BPD total [†]	-0.58 [*]	-0.11	0.06	-0.14	0.10
Affective [†]	-0.46 [*]	-0.11	0.05	-0.08	0.02
Cognitive [†]	-0.47 [*]	-0.08	0.06	-0.07	0.10
Disturbed [†]	-0.35 [*]	-0.11	0.00	-0.11	0.09
Impulsivity [†]	-0.49 [*]	-0.03	0.07	-0.18	0.10
BDI total [†]	-0.04	0.06	-0.06	0.07	0.07
BEST total [†]	-0.10	-0.24 [*]	-0.03	-0.02	0.06
BIS total [†]	-0.04	-0.06	-0.07	0.08	0.18 [*]
SAS total [†]	-0.11	0.00	0.12	0.03	-0.03
SCL-90-R Global Severity Index [†]	-0.02	0.00	0.04	0.11	0.09
Age	-0.11	-0.12	-0.06	-0.07	0.10
Education level achieved	0.01	-0.01	-0.01	-0.12	0.05
Distance traveled to clinic	0.11	-0.01	-0.04	0.11	-0.01
Out-patient mental health visits	-0.03	-0.10	-0.07	0.05	-0.05
No. psychotropic medications taken	0.08	-0.12	-0.12	-0.08	0.17 [*]
No. STEPPS sessions attended	0.09	-0.20	-0.16	0.01	0.50 [*]
No. personality disorders	0.19 [*]	0.01	-0.06	0.01	-0.03
Dichotomous variables					
Treatment group (STEPPS plus TAU)	0.19 [*]	0.16 [*]		0.17 [*] 0.20 [*] 0.12	
Gender (male)	0.13	-0.06	-0.08	0.09	0.02
Individual psychotherapy participation	0.03	0.05	0.06	0.02	0.00
Past psychiatric history	0.01	-0.07	-0.15	0.05	0.09

Baseline predictor variables	Improvement measures				
	ZAN-BPD	BEST	BDI	GAS	LAT
Past self-harm acts	0.19 [*]	0.07	-0.04	0.11	0.06
Past suicide attempt(s)	0.11	0.05	0.00	0.10	0.08
Current major depression	0.17	0.02	-0.06	0.05	0.04
Current anxiety disorder	0.05	0.01	-0.04	0.07	-0.11
Lifetime anxiety disorder	0.04	-0.04	-0.09	-0.02	-0.03
Lifetime substance use disorder	-0.01	0.19 [*]	0.08	0.07	-0.13
Lifetime eating disorder	0.03	-0.16	-0.07	-0.13	0.04
Antisocial PD	0.10	0.12	0.05	0.10	-0.04
Avoidant PD	0.10	-0.04	0.01	-0.05	0.10
Dependent PD	0.09	0.05	0.02	0.10	-0.09
Narcissistic PD	0.13	0.03	-0.09	-0.08	-0.01
Obsessive-compulsive PD	0.04	-0.10	-0.06	0.02	-0.03
Paranoid PD	0.27 [*]	0.01	-0.16	0.08	0.03

LAT, last assessment time; TAU, treatment as usual; PD, personality disorder; CGI, Clinical Global Impression; ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder; BEST, Borderline Evaluation of Severity Over Time; BDI, Beck Depression Inventory; GAS, Global Assessment Scale.

^{*} P -value < 0.05.

[†] Higher scores imply greater symptom severity (less favorable).

Table 3

Multiple logistic regression predictors of response

Baseline predictor	B	SE	OR	χ^2	P
Intercept	-2.032	0.399		26.00	<0.001
ZAN-BPD affective disturbance [†]	-0.225	0.084	0.80	7.12	0.008
Treatment group (STEPPS plus TAU)	2.160	0.467	8.67	21.43	<0.001
Gender (male)	1.320	0.609	3.74	4.69	0.030
ZAN-BPD affective disturbance × group	NS				
Gender (male) × group	NS				

OR, odds ratio or exp(B); ZAN-BPD, Zanarini Rating Scale for Borderline Personality Disorder.

[†]Higher scores imply greater symptom severity (less favorable).

Table 4

Multiple linear regression predictors of improvement measures and last assessment time

Baseline predictor	B	SE	t	P
Improvement in ZAN-BPD total score				
Intercept	0.128	0.092	1.39	0.166
CGI severity [†]	0.050	0.063	0.80	0.428
ZAN-BPD total [†]	-0.062	0.009	-6.87	<0.001
No. personality disorders	-0.037	0.032	-1.13	0.260
Treatment group (STEPPS plus TAU)	0.245	0.081	3.02	0.003
Past self-harm	0.135	0.092	1.47	0.145
CGI severity × treatment group	NS			
ZAN-BPD total × treatment group	0.030	0.012	2.44	0.016
No. personality disorders × treatment group	0.097	0.048	2.03	0.045
Past self-harm, any × treatment group	NS			
Improvement in BEST Total Score				
Intercept	0.010	0.127	0.08	0.938
BEST total [†]	-0.016	0.006	-2.58	0.011
Any substance use disorder	0.208	0.129	1.61	0.109
Treatment group (STEPPS plus TAU)	0.255	0.124	2.05	0.043
BEST total × treatment group	NS			
Any substance use disorder × treatment group	NS			
Improvement in GAS Score				
Intercept	0.146	0.073	2.00	0.048
GAS score	-0.018	0.004	-4.02	<0.001
Treatment group (STEPPS plus TAU)	0.244	0.100	2.43	0.017
GAS score × treatment group	NS			
Last assessment time				
Intercept	33.911	3.268	10.38	<0.001
BIS total [†]	0.462	0.173	2.67	0.009
No. psychotropic medications	1.110	0.873	1.27	0.206
Treatment group (STEPPS plus TAU)	7.666	4.465	1.72	0.088
BIS total × treatment group	NS			
No. psychotropic medications × treatment group	NS			

PD, personality disorder; CGI, Clinical Global Impression; ZAN-BPD, Zanzarini Rating Scale for Borderline Personality Disorder; BEST, Borderline Evaluation of Severity Over Time; BDI, Beck Depression Inventory; GAS, Global Assessment Scale; BIS, Barratt Impulsiveness Scale.

[†] Higher scores imply greater symptom severity (less favorable).