BMJ Open

Emotion Regulation Group Therapy for Deliberate Self-Harm: A Multi-Site Effectiveness Open Trial

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016220
Article Type:	Research
Date Submitted by the Author:	01-Feb-2017
Complete List of Authors:	Sahlin, Hanna; Karolinska Institutet, Department of Clinical Neuroscience, Centre for Psychiatry Research Bjureberg, Johan; Karolinska Institutet, Department of Clinical Neuroscience, Centre for Psychiatry Research Gratz, Kim; University of Toledo, Department of Psychology Tull, Matthew; University of Toledo, Department of Psychology Hedman, Erik; Karolinska Institutet, Department of Clinical Neuroscience Bjärehed, Jonas; Lunds Universitet, Department of Psychology Jokinen, Jussi; Karolinska Institute, Department of Clinical Neuroscience; Umeå University, Department of Clinical Sciences Lundh, Lars-Gunnar; Lunds Universitet, Department of Psychology Ljótsson, Brjánn; Karolinska Institutet, Department of Clinical Neuroscience, Division of Psychology Hellner Gumpert, Clara; Karolinska Institutet, Stockholm, Sweden, Department of Clinical Neuroscience, Stockholm Center for Psychiatry Research and Education
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Evidence based practice
Keywords:	Borderline personality disorder, Effectiveness study, Group therapy, Emotion regulation, Suicide & self-harm < PSYCHIATRY, Implementation

Word count: 3740

Emotion Regulation Group Therapy for Deliberate Self-Harm: A Multi-Site Effectiveness Open Trial

Hanna Sahlin[1¶*], Johan Bjureberg[1¶], Kim L. Gratz[2], Matthew T. Tull[2], Erik Hedman[1], Jonas Bjärehed[3], Jussi Jokinen[1,4], Lars-Gunnar Lundh[3], Brjánn Ljótsson[1], Clara Hellner Gumpert[1]

[1]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

[2]University of Toledo, Department of Psychology, Toledo, OH, USA

[3]Department of Psychology, Lund University, Lund, Sweden

[4]Department of Clinical Sciences, Psychiatry, Umeå University, Umeå, Sweden

*Corresponding Author: Hanna Sahlin, MSc

Email: Hanna.Sahlin@ki.se

¶These authors contributed equally to this manuscript.

ABSTRACT

Objective: Emotion regulation group therapy (ERGT) has shown promising results in several efficacy trials. However, it has not been evaluated outside a research setting. In order to increase the availability of empirically supported treatments for individuals with borderline personality disorder and deliberate self-harm, an effectiveness study of ERGT was conducted with therapists of different professional backgrounds who had received brief intensive training in ERGT prior to trial onset.

Design: Multi-site effectiveness within group study with assessments at pre-treatment, post-treatment, and six-month follow-up.

Setting: 14 adult outpatient psychiatric clinics across Sweden

Participants: Ninety-five women (mean age = 25.1 years) with borderline personality disorder (both threshold and subthreshold) and repeated self-harm were enrolled in the study. Ninety-three percent of participants completed the post-treatment assessment and 88% completed the follow-up assessment.

Primary and secondary outcome measures: Primary outcome was self-harm frequency as measured with the Deliberate Self-Harm Inventory. Secondary outcomes included emotion dysregulation, other self-destructive behaviors, depression, anxiety, and stress symptoms, and interpersonal and vocational difficulties.

Intervention: ERGT is an adjunctive, 14-week, acceptance-based behavioral group treatment that directly targets both self-harm and its proposed underlying mechanism of emotion dysregulation.

Results: At post-treatment, intent-to-treat analyses revealed significant improvements with moderate effect sizes in self-harm frequency (Cohen's *d*, 0.52, 95% CI, 0.30-0.75, $p<0.001$) and secondary outcomes such as emotion dysregulation, other self-destructive behaviors, and general psychiatric symptomatology. These results were either maintained or further improved upon at six-month follow-up.

Conclusions: ERGT appears to be a feasible, transportable, and effective treatment for deliberate self-harm and other self-destructive behaviours, emotion dysregulation, and psychiatric symptoms when delivered by clinicians in the community.

Trial registration: Clinicaltrials.gov identifier: NCT01986257

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This multi-site effectiveness study suggests that emotion regulation group therapy may be an easily disseminated and effective treatment for deliberate self-harm.
- Results revealed continued reductions of deliberate self-harm after treatment conclusion.
- This study lacked a control group, limiting our ability to draw conclusions about the effect of ERGT specifically.
- As only adult women were included in the study, generalizability to other patient populations is unclear.

INTRODUCTION

Deliberate self-harm (DSH; also referred to as nonsuicidal self-injury) is defined as "the deliberate, direct destruction or alteration of body tissue without conscious suicidal intent, but resulting in injury severe enough for tissue damage (e.g., scarring) to occur" ([1], p. 255). DSH is highly prevalent in both clinical and nonclinical adult and adolescent populations,[2-4] and is particularly common among individuals with borderline personality disorder (BPD;[5,6]). DSH has been implicated in the high levels of health care utilization among individuals with BPD[7] and is one of the strongest predictors of future suicide attempts.[8,9]

Although the past two decades have seen the development of several efficacious treatments for DSH within BPD, including dialectical behaviour therapy (DBT;[6]), mentalization-based treatment,[10] and emotion regulation group therapy (ERGT;[11,12]), few studies have examined the effectiveness of these treatments in traditional clinical settings (for exceptions, see[13-16]) and the extent to which they can be disseminated to community clinicians remains unclear ([17]). Indeed, although efficacy trials are imperative for establishing the evidence-base of a treatment, they often maximize internal validity, specialized training, and experimental controls at the cost of external validity and applicability to traditional clinical settings.[18,19] Effectiveness trials, on the other hand, are designed to evaluate how a treatment works under more "real-world" conditions and, as such, contribute important information about the utility and feasibility of a treatment in traditional clinical settings. These trials also play an important role in increasing the availability of evidence-based treatments.[20] Thus, further research examining the effectiveness of empirically supported treatments for DSH within BPD is needed.

Thus, this study sought to examine the effectiveness of one such treatment in a nationwide open trial. Specifically, in an effort to increase the availability of clinically-feasible treatments for DSH in BPD within the community, we evaluated the effectiveness of ERGT (a 14-week adjunctive group treatment with established efficacy in the treatment of DSH within BPD;[12]) as delivered by community clinicians at 14 psychiatric outpatient clinics throughout Sweden. Consistent with past research on ERGT, we expected to find significant improvements from pre- to post-treatment in DSH and other self-destructive behaviours, emotion dysregulation, psychiatric symptoms, and adaptive functioning, as well as stability of these improvements during the 6-month follow-up period.

MATERIALS AND METHODS

Design and Participants

The present effectiveness trial was conducted at 14 psychiatric outpatient clinics located throughout Sweden. We used an open trial design with a six-month follow-up.

Participants were recruited and assessed by community-based health care professionals at the psychiatric outpatient clinics. To ensure comparability of the findings to previous ERGT trials, the inclusion criteria in this study were similar to those used in earlier ERGT studies.[11,12,22] Eligibility criteria included: (a) being a woman ≥18 years; (b) meeting ≥3 diagnostic criteria for BPD as determined by the Structured Clinical Interview for DSM-IV Personality Disorders (SCID-II;[23]); (c) ≥3 episodes of DSH in the past six months as assessed by a clinician-administered interview version of the Deliberate Self-harm Inventory (DSHI;[1]); (d) ongoing

treatment as usual in the community; and (e) stability of psychotropic medications for at least two months before inclusion. Exclusion criteria were minimal and included only: (a) a DSM-IV[24] diagnosis of psychotic or bipolar I disorder or ongoing (past month) substance dependence as assessed with the MINI International Neuropsychiatric Interview (MINI 6;[25]); (b) the presence of comorbid psychiatric disorders that required immediate treatment (e.g., anorexia nervosa); (c) insufficient understanding of the Swedish language; and (d) current life circumstances that would interfere with treatment (i.e., being homeless). The study was approved by the Regional Ethical Review Board in Stockholm (Dnr 2013/1321-31/3) and was registered on Clinicaltrials.gov (Identifier NCT01986257).

Between October 2013 and March 2014, 108 female participants with threshold or subthreshold BPD were considered for participation in the study. All participants provided written informed consent. Eight participants did not meet inclusion criteria; four dropped out before completing the pre-treatment assessment; and one completed the pre-treatment assessment but died from suicide before beginning ERGT. Thus, the final sample size was 95 participants. Diagnostic and demographic data for the final sample are presented in Table 1. Notably, this sample was comparable to those of both past ERGT trials and other BPD treatment outcome studies with regard to both demographic characteristics and co-occurring psychiatric disorders [12,15,22,26,27]. Participant flow through the trial is described in Figure 1.

INSERT FIGURE 1 ABOUT HERE: **Figure 1**. Participant flow through the study

Table 1. Sociodemographic, clinical, and diagnostic data of the sample (N = 95)

Variable	n	(%)	Mean	SD	Range

Age			25.1	7.0	18-49
Educational level					
Primary school	27	(28.4)			
High school/vocational school	58	(61.1)			
University	10	(10.5)			
Marital status					
Single	26	(27.3)			
Married/Cohabiting	24	(25.3)			
Living with children	16	(16.8)			
Occupational status					
Full-time student	37	(39.0)			
Employed	24	(25.3)			
Unemployed	21	(22.1)			
On disability pension	13	(13.7)			
On temporary sick-leave	24	(25.3)			
Clinical characteristics					
Meeting full diagnostic criteria for BPD	65	(68.0)			
Number of threshold BPD criteria			5.2	1.5	
Suicide attempt, lifetime[a]	54	(58.1)			1-50
Suicide attempt, past 3 months[a]	18	(17.2)			1-4
DSH frequency past 6 months			61.4	83.3	
Rate of psychiatric medication use	68	(71.6)			
Number of psychiatric medications			1.9	1.8	
Rate of previous psychiatric treatment	80	(84.2)			
Months of ongoing treatment			12.0	34.6	
Type of ongoing treatment					
Cognitive behavioural therapy	30	(31.6)			
Psychodynamic therapy	10	(10.5)			
Supportive therapy	45	(47.4)			
Other	10	(10.5)			
Co-occurring psychiatric disorders					
Depression	51	(53.7)			
Panic disorder	36	(37.9)			
Social anxiety disorder	36	(37.9)			
Posttraumatic stress disorder	20	(21.1)			
Generalized anxiety disorder	38	(40)			
Eating disorder	15	(15.8)			
Substance use disorders	4	(4.2)			

Note. BPD = Borderline personality disorder, DSH = Deliberate self-harm

[a]There were data missing for two participants on history of attempted suicide (n = 93)

Selection of Participating Clinics and Study Therapists

An invitation to participate in the study was distributed through a national network of psychiatric caregivers (with representatives from all county councils). Thirty-two clinics responded to the

invitation, from which 15 clinics were selected. All participating clinics had to have at least one employed therapist meeting the following criteria: (a) educated within a regulated profession (e.g., a licensed nurse, psychologist, or psychotherapist); and (b) basic training in cognitive behavioural therapy (with training in acceptance and commitment therapy [ACT] and/or DBT preferred). Consideration was also given to the clinics' geographical location, aiming for as broad a national geographical representation as possible. One included clinic did not participate in the study due to local administrative difficulties, leaving 14 clinics and 28 therapists in the study. Across these clinics, a total of 17 groups were conducted. Median number of participants treated at each clinic was seven (IQR: 5-9; min 4, max 11). The clinics were located in 11 cities (population size range: 33 155 – 2 136 042, median 101 615.5;[28]).

To compensate for the extra cost of implementing a new treatment within the context of regular care, the clinics received monetary compensation for administration of ERGT at a value of 1800 USD per group.

Assessments

Clinician-administered assessments at baseline included the BPD module of the SCID-II,[23] an interview version of the DSHI,[1] MINI 6,[25] and the Columbia—Suicide Severity Rating Scale (C-SSRS;[29]). Treatment outcome measures were administered in self-report format at baseline, pre-treatment, post- treatment, and six-month follow-up. All self-report measures used in the study were completed online (a method with demonstrated validity;[30]).

The primary outcome measure was the total frequency of DSH measured by the Deliberate Self-Harm Inventory, a self-report measure with adequate test-retest reliability and construct, discriminant, and convergent validity (DSHI;[1]). The DSHI specifies 16 different types of DSH (e.g., cutting, burning or hitting oneself). The DSHI was also used to assess DSH versatility (i.e., number of different types of DSH behaviours in the past 4 months) – an index of DSH severity.[31]

The secondary outcome measures included the following self-report measures: the Difficulties in Emotion Regulation Scale (DERS;[32]), a measure of clinically relevant emotion regulation difficulties (α=.90 in this sample) with good test-retest reliability and construct and predictive validity;[22,32,33] the 11-item self-destructive behaviour supplement to the Borderline Symptom List (BSL;[34]), which assesses past week engagement in several self-destructive behaviours (e.g., binge eating, excessive drinking, drug-use, risky sexual behaviours); the 21-item Depression Anxiety Stress Scales (DASS-21;[35]), a measure of depression, anxiety, and stress symptoms (α=.88 in this sample) with good test-retest reliability and construct and discriminant validity;[35,36] the BPD-related composite of the Inventory of Interpersonal Problems (IIP-BPD;[37]), a measure of BPD-relevant interpersonal difficulties (α=.89 in this sample) with good convergent validity and specificity;[37] and the Sheehan Disability Scales (SDS;[38]), a widely used measure of social and vocational impairment due to psychological symptoms (α=.80 in this sample) with adequate reliability and construct, convergent, and discriminant validity across various clinical populations.[39,40] Treatment credibility and expectancy were assessed after the second session of ERGT with the Credibility/Expectancy Questionnaire.[41]

Treatment

ERGT is a 14-session, adjunctive, acceptance-based behavioural group treatment developed to treat DSH by targeting its underlying mechanism of emotion dysregulation.[11] Based on Gratz and Roemer's[32] conceptualization of emotion regulation, ERGT systematically teaches skills aimed at improving a number of dimensions of emotion regulation, including: emotional awareness, understanding, and acceptance; the ability to control behaviours when experiencing negative emotions; the use of non-avoidant emotion regulation strategies to modulate the intensity and/or duration of emotional responses; and the willingness to experience negative emotions as part of pursuing meaningful activities in life. Moreover, the following themes are emphasized throughout the treatment: (a) the potentially paradoxical effects of emotional avoidance, (b) the emotion regulating consequences of emotional acceptance and willingness, and (c) the importance of controlling behaviour when emotions are present, rather than controlling emotions themselves. A detailed description of the content and development of ERGT is available elsewhere.[11]

The ERGT treatment manual (Gratz & Tull, 2010, unpublished manual) provides thorough instructions on the therapeutic stance and theoretical framework of ERGT, as well as detailed descriptions of the content, in-session exercises, and homework assignments for each session. Sudden deterioration or suicidal crises are monitored weekly through self-report measures assessing DSH frequency and emotion dysregulation, and addressed by the group therapists in collaboration with the ongoing treatment provider if needed. The ERGT treatment manual was translated into Swedish through a collaborative and iterative process involving the treatment

developer (KLG) and co-author of the ERGT manual (MTT) and the primary ERGT supervisors for this trial (HS and JB, in collaboration with LGL). In this trial, ERGT was delivered in weekly two-hour sessions led by two therapists with groups of 4-9 participants (average group size = 5.59).

Therapist Training and Treatment Fidelity

Prior to the trial, therapists were required to study the ERGT manual and relevant articles on the theoretical underpinnings of ERGT, as well as to participate in a three-day workshop led by the authors of the ERGT manual (KLG and MTT). The workshop consisted of didactic lectures, demonstrations, role-playing, and practice exercises. Of the 28 therapists, 22 were licensed psychologists, two were social workers, two were nurses, one was a psychiatric aid, and one was an occupational therapist. Twenty-six therapists had previous experience treating patients with BPD and/or DSH; three therapists had previous experience with ACT, and 19 therapists had experience with DBT. To ensure treatment fidelity during the ERGT trial, all sessions were filmed and reviewed weekly by clinicians with experience delivering ERGT (HS, JB), and all therapists received the option of weekly supervision based on the reviewed films.

Statistics and Data Analysis

We expected to recruit 90-105 participants. This sample size yielded high power (> .99 with $\alpha = .05$) to detect a standardized mean difference between pre- and post-treatment of $d = 0.5$ using within-group t-tests on a log-transformation of the primary outcome measure DSH-frequency (consistent with past research on ERGT;[11,12,22]).

All analyses were performed in R using random effects modelling [42]. The count variables, DSH frequency and BSL, were analysed using Poisson models, and the remaining continuous outcomes were analysed using linear models. The models included all available data at the three assessment points (pre-treatment, post-treatment, and six-month follow-up) for each outcome, thus making them intent-to-treat analyses. We estimated separate slopes for the change between the pre- and post-treatment assessments (S1) and the change between the post-treatment and six-month follow-up assessments (S2). Random intercepts and random slopes were included in the models if they significantly improved model fit according to log-likelihood ratio tests.

Effect sizes (i.e., Cohen's *d*) for the changes between the three assessments were calculated by dividing the appropriate slope estimate (i.e., pre- to post-treatment: S1, pre-treatment to six-month follow-up: S1+S2, and post-treatment to six-month follow-up: S2) by the pre-treatment standard deviation. We also performed separate analyses where linear mixed models were applied to log-transformed DSH frequency and BSL scores and corresponding effect sizes were extracted (to permit comparison with previous studies of ERGT). However, inferences of the statistical significance of changes on these measures were based on the more appropriate Poisson regression models. Confidence intervals with a 95% margin for the effect sizes were calculated using 5000 bootstrap replications.[43] The bootstrap replications were clustered on participants.[44]

We also performed sensitivity analyses to investigate the robustness and validity of the DSHI results. First, the treating clinics were entered as random factors to test the possibility of

clustering effects in the data. Second, we added concurrent medication status (coded as 0 for no concurrent medication and 1 for concurrent medication) and type of treatment as usual (coded as a factor with the following levels: cognitive behavioural therapy, psychodynamic therapy, supportive therapy, or other) as covariates in the model. These covariates were added both as simple effects and as interaction effects with the S1 variable to investigate the possible influence of the covariates on DSH-frequency and treatment effect. Third, we entered the number of treatment sessions attended as a predictor of improvement in DSH frequency during the treatment period.

Finally, we examined the number of participants who reported no (zero) DSH episodes at each assessment point and used McNemar's exact tests to analyse the changes between the assessment points.

RESULTS

Treatment Adherence and Attrition

The average time from baseline assessment to the start of treatment was 20.7 days (SD = 17.2, range 1-91). Twenty-one participants (22%) dropped out of ERGT (see Figure 1). Mean number of sessions attended for all included participants was 11 (SD = 5.2, min 0, max 16). Seventy-two participants (76%) attended ≥ 7 sessions and 47 (49%) attended 14 sessions. Post-treatment assessments were completed by 88 (93%) participants and 6-month follow-up assessments were completed by 76 (82%) participants. Mean ratings of treatment credibility and expectancy completed after the second session were 5.7 (SD = 1.8) and 47.0% (SD = 25.5), respectively (for comparison, mean ratings reported by Gratz & Tull were 6.91 and 57%, respectively).[22]

Primary Outcome

Results for both the primary and secondary outcome measures are displayed in Table 2. There were significant reductions in DSH frequency and versatility from pre- to post-treatment with moderate effect sizes. Moreover, results revealed further significant improvements in DSH frequency and versatility from post-treatment to the six-month follow-up. The improvements in DSH from pre-treatment to 6-month follow-up were accompanied by medium (for DSH versatility) to large (for DSH frequency) effect sizes.

The observed medians for DSH frequency were 22.0 (IQR: 9.5-56.0), 10.0 (IQR: 2.8-45.5), and 4.0 (IQR: 0.0-13.0) at pre-treatment, post-treatment, and follow-up, respectively. The percentage of participants with an observed score of zero DSH episodes during the past four months increased significantly from 4.2% at pre-treatment to 17.9% at post-treatment (Exact McNemar's $\chi^2 = 12.25$, df = 1, $p < .001$) and to 25.3% at follow-up (Exact McNemar's $\chi^2 = 16.67$, df = 1, $p < .001$). The increase between post-treatment and follow-up was also significant (Exact McNemar's $\chi^2 = 7.14$, df = 1, $p = .01$).

Secondary Outcomes

There were significant improvements in emotion dysregulation, self-destructive behaviours, and depression and stress symptoms at post-treatment, accompanied by small to large effect sizes (see Table 2). At the six-month follow-up, all of these improvements were either maintained or further improved upon, with the change on the DERS from pre-treatment to 6-month follow-up reaching a large-sized effect. The observed median self-destructive behaviour scores were 4.0

(IQR: 2.0-7.0), 2.0 (IQR: 1.0-5.0) and 2.0 (IQR: 0.3-5.0) at pre-treatment, post-treatment, and follow-up, respectively. Improvements in interpersonal difficulties, anxiety, and social and vocational impairment were not significant at post-treatment; however, there were significant improvements in these outcomes (accompanied by small to moderate effect sizes) from pre-treatment to the six-month follow-up stemming from the additional significant improvements in these measures observed from post-treatment to six-month follow-up.

Sensitivity analyses

First, we included the treating clinics as random factors in the DSH frequency model. This did not improve model fit, suggesting that there was not a significant clustering effect of treating clinic in the data. Second, we included concurrent medication status and type of treatment as usual as simple effects and interaction effects with S1 (i.e., the change between the pre-treatment and post-treatment assessments) in the DSH frequency model. None of the added predictors were statistically significant ($p > .05$), suggesting that concurrent medication use and type of treatment as usual were not associated with treatment effect. Third, we included the number of attended sessions as a predictor of improvement in DSH frequency. The interaction effect between session count and S1 was statistically significant, ($B = 0.06$, $Z = 11.54$, $p < .001$), indicating that higher attendance was associated with larger improvements during treatment.

Table 2. Treatment outcome variables at pre-treatment, post-treatment and six-month follow-up

Outcome	Pre-treatment	Post-treatment	6-mo f-u	Pre-to post-treatment comparison		Post- to 6-mo follow-up comparison		Pre-to 6-mo follow-up comparison	
	Mean (SD)	*Mean (SD)*	*Mean (SD)*	*Z*	*Cohen's d [95% CI]*	*Z*	*Cohen's d [95% CI]*	*Z*	*Cohen's d [95% CI]*
DSHI-f	53.68 (99.88)	37.45 (72.22)	28.69 (95.44)	-5.61***		-9.54***			
DSHI-f[a]	3.10 (1.39)	2.39 (1.69)	1.71 (1.62)		0.52 [0.30, 0.75]		0.47 [0.27, 0.70]		0.99 [0.70, 1.30]
DSHI-v	3.01 (1.82)	2.23 (2.02)	1.67 (1.76)	-3.67***	0.41 [0.19, 0.63]	-2.86**	0.24 [0.07, 0.42]	-5.65***	0.65 [0.40, 0.89]
BSL	4.82 (3.69)	3.65 (4.24)	3.24 (3.61)	-5.06***		-0.71			
BSL[a]	1.52 (0.76)	1.19 (0.84)	1.12 (0.83)		0.43 [0.20, 0.65]		0.08 [-0.14, 0.31]		0.51 [0.27, 0.77]
DERS	125.98 (19.37)	108.17 (27.52)	104.66 (27.40)	-6.56***	0.91 [0.63, 1.20]	-1.00	0.12 [-0.09, 0.43]	-7.13***	1.03 [0.69, 1.38]
IIP-BPD	2.05 (0.72)	1.98 (0.71)	1.78 (0.76)	-1.08	0.10 [-0.06, 0.29]	-2.43*	0.24 [0.06, 0.43]	-3.46***	0.34 [0.13, 0.59]
DASS-D	25.35 (10.28)	20.11 (11.80)	19.95 (11.92)	-4.16***	0.50 [0.29, 0.76]	-0.42	0.05 [-0.21, 0.30]	-4.35***	0.56 [0.27, 0.86]
DASS-A	17.14 (8.98)	16.30 (9.97)	14.54 (9.58)	-0.81	0.08 [-0.10, 0.27]	-1.54	0.17 [-0.05, 0.42]	-2.33*	0.25 [0.00, 0.49]
DASS-S	25.77 (7.95)	23.34 (9.21)	21.19 (10.34)	-2.46*	0.30 [0.06, 0.54]	-1.95*	0.26 [0.01, 0.52]	-4.30***	0.56 [0.26, 0.86]
SDS	18.44 (6.99)	18.59 (7.05)	16.01 (8.41)	0.25	-0.03 [-0.22, 0.16]	-2.71**	0.32 [0.07, 0.60]	-2.51*	0.29 [0.04, 0.59]

Note. Test statistics are based on Poisson regression analyses for count data and mixed models analyses for continuous data. Confidence intervals for effect sizes are based on 5000 bootstrap replications. Abbreviations: BSL = Borderline Symptom List, behaviour supplement, DASS-D = Depression Anxiety and Stress

Scales – 21 Depression, DASS-A = Depression Anxiety and Stress Scales – 21 Anxiety, DASS-S = Depression Anxiety and Stress Scales – 21 Stress, DERS = Difficulties in Emotion Regulation Scale, DSHI-f = Deliberate Self-Harm Inventory - frequency, DSHI-v = Deliberate Self-Harm Inventory - versatility, IIP-BPD = BPD-related composite of the Inventory of Interpersonal Problems, SDS = Sheehan Disability Scales.

[a] Effect size estimates for DSHI frequency and BSL scores were based on log-transformed data that were analysed in mixed models analyses.

*p<.05, **p<.01, ***p<.001

DISCUSSION

The present multi-site effectiveness study provides additional support for the feasibility and transportability of ERGT. Results revealed significant improvements in DSH frequency and versatility, emotion dysregulation, self-destructive behaviours, and depression and stress symptoms from pre- to post-treatment. By the six-month follow-up, interpersonal difficulties and social and vocational impairment had also improved significantly. Moreover, all gains found at post-treatment were either maintained or further improved upon at follow-up. Notably, the results of the sensitivity analyses strengthen our confidence in these findings, providing evidence of a significant effect of session attendance, but not treatment clinic, concurrent medication use, or type of treatment as usual, on improvements in DSH frequency. These results are consistent with past findings that characteristics of participants' ongoing therapy in the community had minimal impact on treatment response to ERGT,[20] and suggest that it is engagement in ERGT rather than other (non-specific) treatment-related factors that influences reductions in DSH frequency.

Results from this effectiveness trial are similar to those obtained in previous ERGT efficacy trials,[11,12] which have revealed positive effects of ERGT on DSH, emotion dysregulation, and psychiatric symptoms. Likewise, our findings of either stability or further improvements during the follow-up period are consistent with the observed pattern of continued or maintained improvement during a 9-month follow-up period in Gratz and colleagues' RCT.[12] These findings provide further support for the durability of improvements following this relatively brief

and non-intensive treatment, suggesting that durable gains may be obtained with ERGT even when delivered by community clinicians with only brief training in this treatment.

Despite these similarities with previous ERGT trials, the pattern of findings for measures of interpersonal, social, and vocational functioning differed from Gratz and colleagues' previous ERGT studies ([12,22]). Specifically, results revealed no significant changes in interpersonal difficulties or social and vocational impairment from pre- to post-treatment, although there were significant (albeit small) improvements in these areas during the follow-up period. Although the differences in these findings may be due to study-related differences in therapist training and/or treatment delivery, they may also be explained (in part) by the fact that more than one-third of participants (39%) in the present trial were on disability pension or current sick leave, thus limiting their social interactions and vocational opportunities during treatment.

Notably, rates of abstinence from DSH increased significantly from pre- to post-treatment, as well as from post-treatment through the 6-month follow-up, with 25.2% of participants reporting abstinence from DSH 6-months post-treatment. Nonetheless, it warrants mention that the proportion of participants reaching abstinence from DSH was lower than in previous ERGT trials.[12,22] This may be due to differences in the level of training and supervision provided in this trial versus previous trials. Indeed, ratings of treatment credibility and expectancy in this trial were somewhat lower than in previous studies of ERGT, potentially capturing the lesser experience of the therapists in this trial. Conversely, the lower rate of treatment-related abstinence from DSH observed in this study may reflect differences in sample composition and/or clinical severity (as both emotion dysregulation and DSH frequency reported in this

sample were at the high end of the range reported in past ERGT studies). Nonetheless, past research examining predictors of treatment response to ERGT found that several indicators of greater severity in domains relevant to ERGT (i.e., baseline emotion dysregulation and BPD criteria, lifetime and recent DSH, and past-year hospitalization and suicide attempts) predicted *better* responses during treatment and follow-up,[20], suggesting that greater severity in certain domains may be associated with better response to this treatment. Future research is needed to clarify the particular patients most likely to benefit from ERGT.

It is also important to note that we found significant reductions across all assessment points in DSH versatility (i.e., a marker of DSH severity associated with future suicide risk).[31] This finding lends support to the potential utility of ERGT in decreasing risk for self-injurious behaviours in general. Indeed, given the high occurrence of DSH and suicidal behaviours [9], as well as evidence that DSH is one of the strongest prospective predictors of future suicide attempts,[45,46] the emphasis within ERGT on reducing DSH could be expected to reduce suicidal risk as well. Nonetheless, in the absence of data in this or previous ERGT trials on suicidal outcomes in particular, the benefits of ERGT for suicidality remain unknown and in need of future investigation.

There are several strengths of this study that are worth noting. First, the naturalistic design of this study permitted inclusion of a large sample of participants who received this ERGT as part of their standard treatment. Second, data attrition was low (8% at post-treatment and 18% at follow-up) and participant dropout was within expected levels when treating individuals with BPD[47] and consistent with previous ERGT trials.[12] Third, contrary to previous ERGT trials,[11,12,22]

the group leaders were of different professional backgrounds and representative of the community clinicians who regularly treat this patient population, providing support for the generalizability and transportability of this treatment. Finally, our data provide support for the feasibility of disseminating ERGT to community clinicians, as therapists in this study were provided with only readings and a brief workshop on ERGT prior to its delivery.

Despite these strengths, there are also some limitations. First, the lack of a control condition precludes conclusions about the effects of this ERGT (vs. treatment as usual or the passage of time). However, it is important to note that the waitlist conditions in previous ERGT trials generally evidence stability over time on all measures of interest (likely due to the relatively short time-frame of the treatment period). Nonetheless, future studies are needed to evaluate the effectiveness of ERGT within an RCT design. Second, even though the participating clinicians were offered weekly supervision based on filmed sessions, no systematic adherence ratings were conducted. This limits our ability to speak to the quality of ERGT and its relation to outcome. Third, the results relied solely on self-report measures and not clinician-rated outcomes. However, research on self-reported versus clinician-rated outcomes in psychotherapy studies suggests that the use of self-report measures often results in smaller effect sizes when compared to clinician administered instruments.[48] Thus, it is possible that our results may be conservative estimates of improvements during and after treatment. Finally, our study only included women, which limits the generalizability of the results to men. Future research is needed that evaluates ERGT within male or mixed-gender samples.

Despite these limitations, our results provide further evidence for the utility and transportability of ERGT, suggesting that this is a feasible and effective treatment for DSH, emotion dysregulation, and psychiatric symptoms when delivered by community clinicians in traditional clinic settings.

ACKNOWLEDGEMENTS

The authors wish to thank Ida Janson, Sara Eivergård, Niclas Andersson and all study therapists for their hard work and enthusiasm, making this study possible.

COMPETING INTERESTS

Ms Hanna Sahlin, Mr Johan Bjureberg, Dr Kim L Gratz, Dr Matthew T Tull, Dr Jonas Bjärehed, Dr Lars-Gunnar Lundh, Dr Jussi Jokinen and Dr Clara Hellner Gumpert report no competing interests. Dr Erik Hedman and Dr Brjánn Ljótsson are shareholders of a company, Dahlia, specialized in psychiatric symptom assessment.

FUNDING

This research was supported by the National Self Harm project in Sweden and Stockholm County Council regional research grant #SLL20140428. None of the funding organizations has had any role in the design and conduct of the study; in the collection, management, and analysis of the data; or in the preparation, review and approval of the manuscript.

DATA SHARING

No additional data are available.

AUTHOR CONTRIBUTIONS

Drs Sahlin, Bjureberg, Ljótsson and Gumpert had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Drs Sahlin, Bjureberg, Gratz, Tull, Hedman, Bjärehed, Lundh, Jokinen, Ljótsson and Gumpert. Drs Sahlin, Bjureberg, Bjärehed, Ljótsson and Gumpert

Drafting of the manuscript: Drs Sahlin, Bjureberg, Ljótsson and Gumpert

Critical revision of the manuscript for important intellectual content and approval of the final version to be published: Drs Sahlin, Bjureberg, Gratz, Tull, Hedman, Bjärehed, Lundh, Jokinen, Ljótsson and Gumpert

TRIAL REGISTRATION

This trial was registered on Clinicaltrials.gov on November 4, 2013.

Clinicaltrials.gov identifier NCT01986257

REFERENCES

1 Gratz K. Measurement of Deliberate Self-Harm: Preliminary Data on the Deliberate Self-Harm Inventory. *Journal of Psychopathol Behav Assess* 2001;**23**:253–63. doi:10.1023/A%3A1012779403943

2 Briere J, Gil E. Self-mutilation in clinical and general population samples: prevalence, correlates, and functions. *Am J Orthopsychiatry* 1998;**68**:609–20.

3 de Klerk S, van Noorden MS, van Giezen AE, *et al.* Prevalence and correlates of lifetime deliberate self-harm and suicidal ideation in naturalistic outpatients: The Leiden Routine Outcome Monitoring study. *J Affect Disord* 2011;**133**:257–64. doi:10.1016/j.jad.2011.03.021

4 Swannell SV, Martin GE, Page A, *et al.* Prevalence of nonsuicidal self-injury in nonclinical samples: systematic review, meta-analysis and meta-regression. *Suicide Life Threat Behav* 2014;**44**:273–303. doi:10.1111/sltb.12070

5 Sansone RA, Wiederman MW, Sansone LA. The Self-Harm Inventory (SHI): development of a scale for identifying self-destructive behaviors and borderline personality disorder. *J Clin Psychol* 1998;**54**:973–83.

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

7 Zanarini MC. Psychotherapy of borderline personality disorder. *Acta Psychiatr Scand* 2009;**120**:373–7. doi:10.1111/j.1600-0447.2009.01448.x

8 Victor SE, Klonsky ED. Correlates of suicide attempts among self-injurers: A meta-analysis. *Clin Psychol Rev* 2014;**34**:282–97. doi:10.1016/j.cpr.2014.03.005

9 Hamza CA, Stewart SL, Willoughby T. Examining the link between nonsuicidal self-injury and suicidal behavior: a review of the literature and an integrated model. *Clin Psychol Rev* 2012;**32**:482–95. doi:10.1016/j.cpr.2012.05.003

10 Bateman A, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *Am J Psychiatry* 1999;**156**:1563–9. doi:10.1176/ajp.156.10.1563

11 Gratz KL, Gunderson JG. Preliminary data on an acceptance-based emotion regulation group intervention for deliberate self-harm among women with borderline personality disorder. *Behav Ther* 2006;**37**:25–35. doi:10.1016/j.beth.2005.03.002

12 Gratz KL, Tull MT, Levy R. Randomized 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. *Psychol Med* 2014;**44**:2099–112. doi:10.1017/S0033291713002134

13 Bales D, van Beek N, Smits M, *et al.* Treatment outcome of 18-month, day hospital mentalization-based treatment (MBT) in patients with severe borderline personality disorder in the Netherlands. *J Pers Disord* 2012;**26**:568–82. doi:10.1521/pedi.2012.26.4.568

14 Feigenbaum JD. A real-world study of the effectiveness of DBT in the UK National Health Service. *Br J Clin Psychol* 2012;**51**:121–41. doi:10.1111/j.2044-8260.2011.02017.x

15 Pasieczny N, Connor J. The effectiveness of dialectical behaviour therapy in routine public mental health settings: an australian controlled trial. *Behav Res Ther* 2011;**49**:4–10. doi:10.1016/j.brat.2010.09.006

16 Wagner T, Fydrich T, Stiglmayr C, *et al.* Societal cost-of-illness in patients with borderline personality disorder one year before, during and after dialectical behavior therapy in routine outpatient care. *Behav Res Ther* 2014;**61**:12–22. doi:10.1016/j.brat.2014.07.004

17 Swales MA, Taylor B, Hibbs RAB. Implementing Dialectical Behaviour Therapy: programme survival in routine healthcare settings. *J Ment Health* 2012;**21**:548–55. doi:10.3109/09638237.2012.689435

18 Weisz JR, Weiss B, Donenberg GR. The lab versus the clinic. Effects of child and adolescent psychotherapy. *Am Psychol* 1992;**47**:1578–85.

19 Wells KB. Treatment Research at the Crossroads: The Scientific Interface of Clinical Trials and Effectiveness Research. *Am J Psychiatry* 1999;**156**:5–10.

20 Gratz KL, Dixon-Gordon KL, Tull MT. Predictors of treatment response to an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Personal Disord* 2014;**5**:97–107. doi:10.1037/per0000062

21 Weiss NH, Tull MT, Viana AG, *et al.* Impulsive behaviors as an emotion regulation strategy: examining associations between PTSD, emotion dysregulation, and impulsive behaviors among substance dependent inpatients. *J Anxiety Disord* 2012;**26**:453–8. doi:10.1016/j.janxdis.2012.01.007

22 Gratz KL, Tull MT. Extending research on the utility of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality pathology. *Personal Disord* 2011;**2**:316–26. doi:10.1037/a0022144

23 First MB, Gibbon M, Spitzer RL, *et al. Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II)*. Washington, D.C: American Psychiatric Press, Inc. 1997.

24 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 4 ed. Washington, DC: Author, 2000.

25 Sheehan DV, Lecrubier Y, Sheehan KH, *et al.* The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;**59 Suppl 20**:22-33.

26 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**:757–66. doi:10.1001/archpsyc.63.7.757

27 McMain SF, Links PS, Gnam WH, *et al.* A Randomized Trial of Dialectical Behavior Therapy Versus General Psychiatric Management for Borderline Personality Disorder. *Am J Psychiatry* 2009;**166**:1365–74. doi:10.1176/appi.ajp.2009.09010039

28 Statistics Sweden. Population in the Country, Counties and Municipalities on 31/12/2013 and Population Change in 2013. 2014. http://www.scb.se/en_/Yearly-statistics--Municipalities-Counties-and-the-whole-country/370301/

29 Posner K. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry*;**168**:1266–77. doi:10.1176/appi.ajp.2011.10111704

30 Hedman E, Ljótsson B, Andersson E, *et al.* Effectiveness and cost offset analysis of group CBT for hypochondriasis delivered in a psychiatric setting: an open trial. *Cogn Behav Ther* 2010;**39**:239–50. doi:10.1080/16506073.2010.496460

31 Turner BJ, Layden BK, Butler SM, *et al.* How often, or how many ways: clarifying the relationship between non-suicidal self-injury and suicidality. *Arch Suicide Res* 2013;**17**:397–415. doi:10.1080/13811118.2013.802660

32 Gratz K, Roemer L. Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure, and Initial Validation of the Difficulties in Emotion Regulation Scale. *J Psychopathol Behav Assess* 2004;**26**:41–54. doi:10.1023/B%3AJOBA.0000007455.08539.94

33 Bjureberg J, Ljótsson B, Tull MT, *et al.* Development and Validation of a Brief Version of the Difficulties in Emotion Regulation Scale: The DERS-16. *J Psychopathol Behav Assess* 2015:1–13. doi:10.1007/s10862-015-9514-x

34 Bohus M, Limberger MF, Frank U, *et al.* [Development of the Borderline Symptom List]. *Psychother Psychosom Med Psychol* 2001;**51**:201–11. doi:10.1055/s-2001-13281

35 Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behav Res Ther* 1995;**33**:335–43.

36 Antony MM, Bieling PJ, Cox BJ, *et al.* Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychol Assess* 1998;**10**:176–81. doi:10.1037/1040-3590.10.2.176

37 Lejuez CW, Daughters SB, Nowak JA, *et al.* Examining the inventory of interpersonal problems as a tool for conducting analogue studies of mechanisms underlying Borderline Personality Disorder. *J Behav Ther Exp Psychiatry* 2003;**34**:313–24. doi:10.1016/j.jbtep.2003.11.002

38 Sheehan DV. *The Sheehan disability scales. The anxiety disease and how to overcome it.* New York City: Charles Scribner and Sons, 1983.

39 Hambrick JP, Turk CL, Heimberg RG, *et al.* Psychometric properties of disability measures among patients with social anxiety disorder. *J Anxiety Disord* 2004;**18**:825–39. doi:10.1016/j.janxdis.2003.10.004

40 Diefenbach GJ, Abramowitz JS, Norberg MM, *et al.* Changes in quality of life following cognitive-behavioral therapy for obsessive-compulsive disorder. *Behav Res Ther* 2007;**45**:3060–8. doi:10.1016/j.brat.2007.04.014

41 Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry* 2000;**31**:73–86.

42 R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: 2015. http://www.R-project.org/

43 Kelley K. The Effects of Nonnormal Distributions on Confidence Intervals Around the Standardized Mean Difference: Bootstrap and Parametric Confidence Intervals. *Educational and Psychological Measurement* 2005;**65**:51–69. doi:10.1177/0013164404264850

44 Ren S, Lai H, Tong W, *et al.* Nonparametric bootstrapping for hierarchical data. *Journal of Applied Statistics* 2010;**37**:1487–98. doi:10.1080/02664760903046102

45 Asarnow JR, Porta G, Spirito A, *et al.* Suicide attempts and nonsuicidal self-injury in the treatment of resistant depression in adolescents: findings from the TORDIA study. *J Am Acad Child Adolesc Psychiatry* 2011;**50**:772–81. doi:10.1016/j.jaac.2011.04.003

46 Wilkinson P, Kelvin R, Roberts C, *et al.* Clinical and psychosocial predictors of suicide attempts and nonsuicidal self-injury in the Adolescent Depression Antidepressants and Psychotherapy Trial (ADAPT). *Am J Psychiatry* 2011;**168**:495–501. doi:10.1176/appi.ajp.2010.10050718

47 Barnicot K, Katsakou C, Marougka S, *et al.* Treatment completion in psychotherapy for borderline personality disorder: a systematic review and meta-analysis. *Acta Psychiatr Scand* 2011;**123**:327–38. doi:10.1111/j.1600-0447.2010.01652.x

48 Cuijpers P, Li J, Hofmann SG, *et al.* Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: a meta-analysis. *Clin Psychol Rev* 2010;**30**:768–78. doi:10.1016/j.cpr.2010.06.001

Figure 1. Participant flow through the study

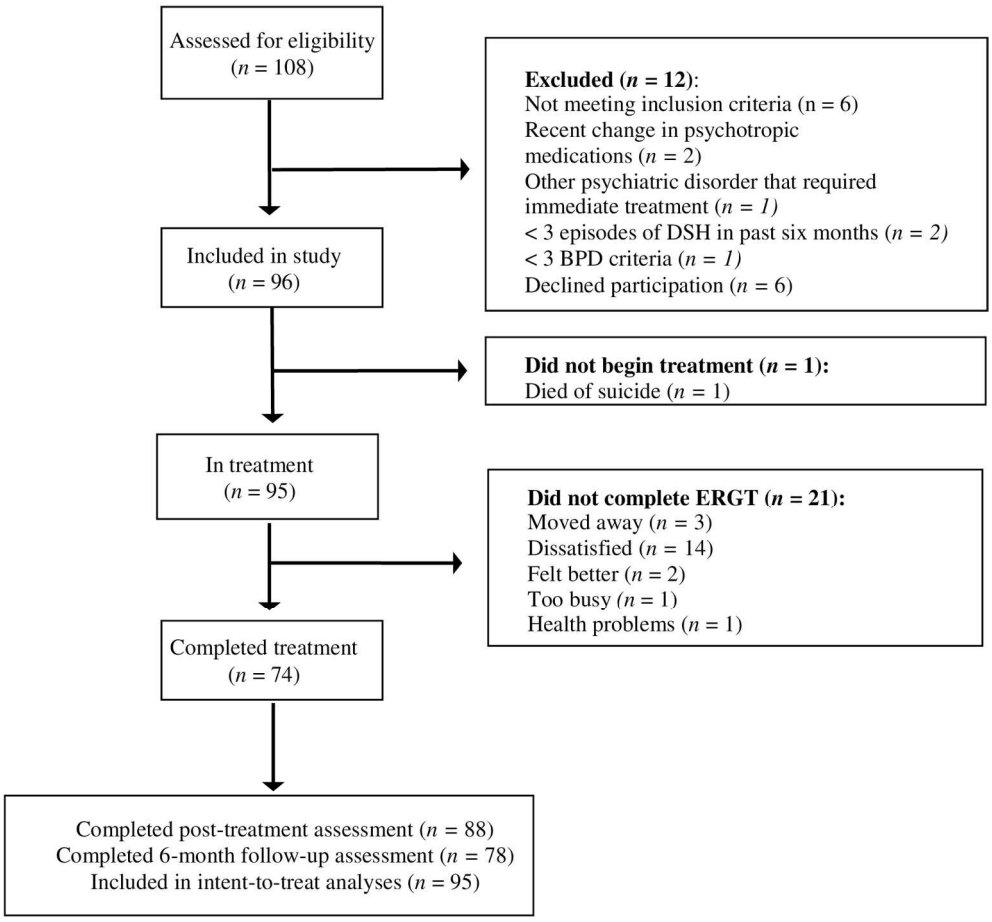

Figure 1. Participant flow through the study

157x162mm (300 x 300 DPI)

TREND Statement Checklist

Paper Section/ Topic	Item No	Descriptor	Reported? ✔	Reported? Pg #
Title and Abstract				
Title and Abstract	1	• Information on how unit were allocated to interventions	✔	2
		• Structured abstract recommended	✔	2
		• Information on target population or study sample	✔	2
Introduction				
Background	2	• Scientific background and explanation of rationale	✔	4
		• Theories used in designing behavioral interventions	✔	4, 9
Methods				
Participants	3	• Eligibility criteria for participants, including criteria at different levels in recruitment/sampling plan (e.g., cities, clinics, subjects)	✔	5
		• Method of recruitment (e.g., referral, self-selection), including the sampling method if a systematic sampling plan was implemented	✔	4
		• Recruitment setting	✔	4-5
		• Settings and locations where the data were collected	✔	6
Interventions	4	• Details of the interventions intended for each study condition and how and when they were actually administered, specifically including:		
		○ Content: what was given?	✔	8
		○ Delivery method: how was the content given?	✔	8
		○ Unit of delivery: how were the subjects grouped during delivery?	✔	7
		○ Deliverer: who delivered the intervention?	✔	7, 9
		○ Setting: where was the intervention delivered?	✔	6
		○ Exposure quantity and duration: how many sessions or episodes or events were intended to be delivered? How long were they intended to last?	✔	8
		○ Time span: how long was it intended to take to deliver the intervention to each unit?	✔	5
		○ Activities to increase compliance or adherence (e.g., incentives)		
Objectives	5	• Specific objectives and hypotheses	✔	5
Outcomes	6	• Clearly defined primary and secondary outcome measures	✔	8
		• Methods used to collect data and any methods used to enhance the quality of measurements	✔	8
		• Information on validated instruments such as psychometric and biometric properties	✔	9
Sample Size	7	• How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules	✔	10
Assignment Method	8	• Unit of assignment (the unit being assigned to study condition, e.g., individual, group, community)	✔	7, 9
		• Method used to assign units to study conditions, including details of any restriction (e.g., blocking, stratification, minimization)		
		• Inclusion of aspects employed to help minimize potential bias induced due to non-randomization (e.g., matching)		

TREND Statement Checklist

Blinding (masking)	9	• Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to study condition assignment; if so, statement regarding how the blinding was accomplished and how it was assessed.		
Unit of Analysis	10	• Description of the smallest unit that is being analyzed to assess intervention effects (e.g., individual, group, or community)	✔	11
		• If the unit of analysis differs from the unit of assignment, the analytical method used to account for this (e.g., adjusting the standard error estimates by the design effect or using multilevel analysis)	✔	10
Statistical Methods	11	• Statistical methods used to compare study groups for primary methods outcome(s), including complex methods of correlated data	✔	10
		• Statistical methods used for additional analyses, such as a subgroup analyses and adjusted analysis	✔	10
		• Methods for imputing missing data, if used	✔	10
		• Statistical software or programs used	✔	10
Results				
Participant flow	12	• Flow of participants through each stage of the study: enrollment, assignment, allocation, and intervention exposure, follow-up, analysis (a diagram is strongly recommended)	✔	21, Figure 1
		○ Enrollment: the numbers of participants screened for eligibility, found to be eligible or not eligible, declined to be enrolled, and enrolled in the study	✔	21
		○ Assignment: the numbers of participants assigned to a study condition	✔	21
		○ Allocation and intervention exposure: the number of participants assigned to each study condition and the number of participants who received each intervention	✔	21
		○ Follow-up: the number of participants who completed the follow-up or did not complete the follow-up (i.e., lost to follow-up), by study condition	✔	21, 11
		○ Analysis: the number of participants included in or excluded from the main analysis, by study condition	✔	21, 11
		• Description of protocol deviations from study as planned, along with reasons		
Recruitment	13	• Dates defining the periods of recruitment and follow-up	✔	5
Baseline Data	14	• Baseline demographic and clinical characteristics of participants in each study condition	✔	6
		• Baseline characteristics for each study condition relevant to specific disease prevention research		
		• Baseline comparisons of those lost to follow-up and those retained, overall and by study condition	✔	10
		• Comparison between study population at baseline and target population of interest	✔	16-17
Baseline equivalence	15	• Data on study group equivalence at baseline and statistical methods used to control for baseline differences		

TREND Statement Checklist

Numbers analyzed	16	• Number of participants (denominator) included in each analysis for each study condition, particularly when the denominators change for different outcomes; statement of the results in absolute numbers when feasible	✔	
		• Indication of whether the analysis strategy was "intention to treat" or, if not, description of how non-compliers were treated in the analyses	✔	10
Outcomes and estimation	17	• For each primary and secondary outcome, a summary of results for each estimation study condition, and the estimated effect size and a confidence interval to indicate the precision	✔	14
		• Inclusion of null and negative findings	✔	13
		• Inclusion of results from testing pre-specified causal pathways through which the intervention was intended to operate, if any		
Ancillary analyses	18	• Summary of other analyses performed, including subgroup or restricted analyses, indicating which are pre-specified or exploratory	✔	11
Adverse events	19	• Summary of all important adverse events or unintended effects in each study condition (including summary measures, effect size estimates, and confidence intervals)		
DISCUSSION				
Interpretation	20	• Interpretation of the results, taking into account study hypotheses, sources of potential bias, imprecision of measures, multiplicative analyses, and other limitations or weaknesses of the study	✔	16
		• Discussion of results taking into account the mechanism by which the intervention was intended to work (causal pathways) or alternative mechanisms or explanations		
		• Discussion of the success of and barriers to implementing the intervention, fidelity of implementation	✔	16,17
		• Discussion of research, programmatic, or policy implications	✔	17-18
Generalizability	21	• Generalizability (external validity) of the trial findings, taking into account the study population, the characteristics of the intervention, length of follow-up, incentives, compliance rates, specific sites/settings involved in the study, and other contextual issues	✔	17-18
Overall Evidence	22	• General interpretation of the results in the context of current evidence and current theory	✔	17-18

From: Des Jarlais, D. C., Lyles, C., Crepaz, N., & the Trend Group (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health*, 94, 361-366. For more information, visit: http://www.cdc.gov/trendstatement/

BMJ Open

Emotion Regulation Group Therapy for Deliberate Self-Harm: A Multi-Site Evaluation in Routine Care using an Uncontrolled Open Trial Design

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016220.R1
Article Type:	Research
Date Submitted by the Author:	16-Jun-2017
Complete List of Authors:	Sahlin, Hanna; Karolinska Institutet, Department of Clinical Neuroscience, Centre for Psychiatry Research Bjureberg, Johan; Karolinska Institutet, Department of Clinical Neuroscience, Centre for Psychiatry Research Gratz, Kim; University of Toledo, Department of Psychology Tull, Matthew; University of Toledo, Department of Psychology Hedman, Erik; Karolinska Institutet, Department of Clinical Neuroscience Bjärehed, Jonas; Lunds Universitet, Department of Psychology Jokinen, Jussi; Karolinska Institute, Department of Clinical Neuroscience; Umeå University, Department of Clinical Sciences Lundh, Lars-Gunnar; Lunds Universitet, Department of Psychology Ljótsson, Brjánn; Karolinska Institutet, Department of Clinical Neuroscience, Division of Psychology Hellner Gumpert, Clara; Karolinska Institutet, Stockholm, Sweden, Department of Clinical Neuroscience, Stockholm Center for Psychiatry Research and Education
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Evidence based practice
Keywords:	Borderline personality disorder, Group therapy, Emotion regulation, Suicide & self-harm < PSYCHIATRY, Implementation

Word count: 4326

Emotion Regulation Group Therapy for Deliberate Self-Harm: A Multi-Site Evaluation in Routine Care using an Uncontrolled Open Trial Design

Hanna Sahlin[1¶*], Johan Bjureberg[1¶], Kim L. Gratz[2], Matthew T. Tull[2], Erik Hedman[1], Jonas Bjärehed[3], Jussi Jokinen[1,4], Lars-Gunnar Lundh[3], Brjánn Ljótsson[1], Clara Hellner Gumpert[1]

[1]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

[2]Department of Psychology, University of Toledo, Toledo, OH, USA

[3]Department of Psychology, Lund University, Lund, Sweden

[4]Department of Clinical Sciences, Psychiatry, Umeå University, Umeå, Sweden

*Corresponding Author: Hanna Sahlin, MSc

Email: Hanna.Sahlin@ki.se

[¶]These authors contributed equally to this manuscript.

ABSTRACT

Objective: Emotion regulation group therapy (ERGT) has shown promising results in several efficacy trials. However, it has not been evaluated outside a research setting. In order to increase the availability of empirically supported treatments for individuals with borderline personality disorder and deliberate self-harm, an evaluation of ERGT in routine clinical care was conducted with therapists of different professional backgrounds who had received brief intensive training in ERGT prior to trial onset.

Design: Multi-site evaluation, using an uncontrolled open trial design with assessments at pre-treatment, post-treatment, and six-month follow-up.

Setting: 14 adult outpatient psychiatric clinics across Sweden

Participants: Ninety-five women (mean age = 25.1 years) with borderline personality disorder (both threshold and subthreshold) and repeated self-harm were enrolled in the study. Ninety-three percent of participants completed the post-treatment assessment and 88% completed the follow-up assessment.

Primary and secondary outcome measures: Primary outcome was self-harm frequency as measured with the Deliberate Self-Harm Inventory. Secondary outcomes included self-harm versatility, emotion dysregulation, other self-destructive behaviours, depression, anxiety, stress symptoms, and interpersonal and vocational difficulties.

Intervention: ERGT is an adjunctive, 14-week, acceptance-based behavioural group treatment that directly targets both self-harm and its proposed underlying mechanism of emotion dysregulation.

Results: At post-treatment, intent-to-treat analyses revealed a significant improvement associated with a moderate effect size on the primary outcome of self-harm frequency (51%, reduction; *Cohen's d* = 0.52, $p < .001$) as well as significant improvements in the secondary outcomes of self-harm versatility, emotion dysregulation, other self-destructive behaviours, and general psychiatric symptomatology. These results were either maintained or further improved upon at six-month follow-up.

Conclusions: ERGT appears to be a feasible, transportable, and useful treatment for deliberate self-harm and other self-destructive behaviours, emotion dysregulation, and psychiatric symptoms when delivered by clinicians in the community.

Trial registration: Clinicaltrials.gov identifier: NCT01986257

STRENGTHS AND LIMITATIONS OF THIS STUDY

- This multi-site evaluation in routine clinical care suggests that emotion regulation group therapy may be an easily disseminated and useful treatment for deliberate self-harm.
- Participants reported continued reductions in deliberate self-harm after treatment conclusion.
- This study lacked a control group, limiting our ability to draw conclusions about the effectiveness of ERGT specifically.
- As only adult women were included in the study, generalizability to other patient populations is unclear.

INTRODUCTION

Deliberate self-harm (DSH; also referred to as nonsuicidal self-injury) is defined as "the deliberate, direct destruction or alteration of body tissue without conscious suicidal intent, but resulting in injury severe enough for tissue damage (e.g., scarring) to occur" ([1], p. 255). DSH is highly prevalent in both clinical and nonclinical adult and adolescent populations,[2-4] and is particularly common among individuals with borderline personality disorder (BPD;[5,6]). DSH has been implicated in the high levels of health care utilization among individuals with BPD[7] and is one of the strongest predictors of future suicide attempts.[8,9]

Although the past two decades have seen the development of several efficacious treatments for DSH within BPD, including dialectical behaviour therapy (DBT;[6]), mentalization-based treatment,[10], and emotion regulation group therapy (ERGT;[11,12]), few studies have evaluated these treatments in traditional clinical settings (for exceptions, see[13-16]) and the extent to which they can be disseminated to community clinicians remains unclear.[17] Indeed, although efficacy trials are imperative for establishing the evidence-base of a treatment, they often maximize internal validity, specialized training, and experimental controls at the cost of external validity and applicability to traditional clinical settings.[18,19] Trials of treatments provided in routine clinical settings, on the other hand, can evaluate how a treatment works under more real-world conditions and, as such, contribute important information about the utility and feasibility of a treatment in traditional clinical settings. These trials also play an important role in increasing the availability of evidence-based treatments.[20] Thus, further research

examining the utility of empirically supported treatments for DSH within BPD in real world clinical settings is needed.

This study sought to examine the utility of one such treatment in a nationwide open trial. Specifically, in an effort to increase the availability of clinically-feasible treatments for DSH in BPD within the community, we conducted an evaluation of ERGT (a 14-week adjunctive group treatment with established efficacy in the treatment of DSH within BPD;[12]) in routine clinical care, as delivered by community clinicians at 14 psychiatric outpatient clinics throughout Sweden. Consistent with past research on ERGT, we expected to find significant improvements from pre- to post-treatment in DSH and other self-destructive behaviours, emotion dysregulation, psychiatric symptoms, and adaptive functioning, as well as stability of these improvements during the six-month follow-up period.

MATERIALS AND METHODS

Design and Participants

The present trial was conducted at 14 psychiatric outpatient clinics located throughout Sweden. We used an uncontrolled open trial design with a six-month follow-up.

Participants were recruited and assessed by community-based health care professionals at the psychiatric outpatient clinics. To ensure comparability of the findings to previous ERGT trials, the inclusion criteria in this study were similar to those used in earlier ERGT studies.[11,12,20] Eligibility criteria included: (a) being a woman ≥18 years; (b) meeting ≥3 diagnostic criteria for BPD as determined by the Structured Clinical Interview for DSM-IV Personality Disorders

(SCID-II;[21]); (c) ≥3 episodes of DSH in the past six months as assessed by a clinician-administered interview version of the Deliberate Self-harm Inventory (DSHI;[1]); (d) ongoing treatment as usual in the community; and (e) stability of psychotropic medications for at least two months before inclusion. Exclusion criteria were minimal and included only: (a) a DSM-IV[22] diagnosis of psychotic or bipolar I disorder or ongoing (past month) substance dependence as assessed with the MINI International Neuropsychiatric Interview (MINI 6;[23]); (b) the presence of co-occurring psychiatric disorders that required immediate treatment (e.g., anorexia nervosa); (c) insufficient understanding of the Swedish language; and (d) current life circumstances that would interfere with treatment (i.e., being homeless). The study was approved by the Regional Ethical Review Board in Stockholm (Dnr 2013/1321-31/3) and was registered on Clinicaltrials.gov (Identifier NCT01986257). Due to an administrative error, the study was not released on the Clinical Trials platform by the Principal Investigator, until November 4, four weeks after the first patient was included. At that date, one group had had one ERGT-session and another group had had two ERGT-sessions, but there were no changes in the study protocol at any time between the start of inclusion of patients and the release of the protocol.

Between October 2013 and March 2014, 108 female participants with threshold or subthreshold BPD were considered for participation in the study. All participants provided written informed consent. Eight participants did not meet inclusion criteria; four dropped out before completing the pre-treatment assessment; and one completed the pre-treatment assessment but died from suicide before beginning ERGT. Thus, the final sample size was 95 participants. Diagnostic and demographic data for the final sample are presented in Table 1. Notably, this sample was comparable to those of both past ERGT trials and other BPD treatment outcome

studies with regard to both demographic characteristics and co-occurring psychiatric disorders.[12,15,20,24,25] Participant flow through the trial is described in Figure 1.

INSERT FIGURE 1 ABOUT HERE: **Figure 1**. Participant flow through the study

Table 1. Sociodemographic, clinical, and diagnostic data of the sample (N = 95)

Variable	n	(%)	Mean	SD	Range
Age			25.1	7.0	18-49
Educational level					
Primary school	27	(28.4)			
High school/vocational school	58	(61.1)			
University	10	(10.5)			
Marital status					
Single	26	(27.3)			
Married/Cohabiting	24	(25.3)			
Living with children	16	(16.8)			
Occupational status					
Full-time student	37	(39.0)			
Employed	24	(25.3)			
Unemployed	21	(22.1)			
On disability pension	13	(13.7)			
On temporary sick-leave	24	(25.3)			
Clinical characteristics					
Meeting full diagnostic criteria for BPD	65	(68.0)			
Number of threshold BPD criteria			5.2	1.5	
Suicide attempt, lifetime[a]	54	(58.1)			1-50
Suicide attempt, past 3 months[a]	18	(17.2)			1-4
DSH frequency past 6 months			61.4	83.3	
Rate of psychiatric medication use	68	(71.6)			
Number of psychiatric medications			1.9	1.8	
Rate of previous psychiatric treatment	80	(84.2)			
Months of ongoing treatment			12.0	34.6	
Type of ongoing treatment					
Cognitive behavioural therapy	30	(31.6)			
Psychodynamic therapy	10	(10.5)			
Supportive therapy	45	(47.4)			
Other	10	(10.5)			
Co-occurring psychiatric disorders					
Depression	51	(53.7)			
Panic disorder	36	(37.9)			
Social anxiety disorder	36	(37.9)			
Posttraumatic stress disorder	20	(21.1)			
Generalized anxiety disorder	38	(40)			
Eating disorder	15	(15.8)			

Substance use disorders	4	(4.2)

Note. BPD = Borderline personality disorder, DSH = Deliberate self-harm

[a]There were data missing for two participants on history of attempted suicide (n = 93)

Selection of Participating Clinics and Study Therapists

An invitation to participate in the study was distributed through a national network of psychiatric caregivers (with representatives from all county councils). Thirty-two clinics responded to the invitation, from which 15 clinics were selected. All participating clinics had to have at least one employed therapist meeting the following criteria: (a) educated within a regulated profession (e.g., a licensed nurse, psychologist, or psychotherapist); and (b) basic training in cognitive behavioural therapy (with training in acceptance and commitment therapy [ACT] and/or DBT preferred). Consideration was also given to the clinics' geographical location, aiming for as broad a national geographical representation as possible. One included clinic did not participate in the study due to local administrative difficulties, leaving 14 clinics and 28 therapists in the study. Across these clinics, a total of 17 groups were conducted. Median number of participants treated at each clinic was seven (IQR: 5-9; min 4, max 11). The clinics were located in 11 cities (population size range: 33 155 – 2 136 042, median 101 615.5;[26]).

To compensate for the extra cost of implementing a new treatment within the context of regular care, the clinics received monetary compensation for administration of ERGT at a value of 1800 USD per group.

Assessments

Clinician-administered assessments at baseline included the BPD module of the SCID-II,[21] an interview version of the DSHI,[1] MINI 6,[23] and the Columbia—Suicide Severity Rating Scale (C-SSRS;[27]). Treatment outcome measures were administered in self-report format at baseline, pre-treatment, post- treatment, and six-month follow-up. All self-report measures used in the study were completed online (a method with demonstrated validity;[28]).

The primary outcome measure was the total frequency of DSH measured by the Deliberate Self-Harm Inventory, a self-report measure with adequate test-retest reliability and construct, discriminant, and convergent validity (DSHI;[1]). The DSHI specifies 16 different types of DSH (e.g., cutting, burning, or hitting oneself). The DSHI was also used to assess DSH versatility (i.e., number of different types of DSH behaviours in the past 4 months) – an index of DSH severity.[29]

The secondary outcome measures included the following self-report measures: the Difficulties in Emotion Regulation Scale (DERS;[30]), a measure of clinically relevant emotion regulation difficulties (α=.90 in this sample) with good test-retest reliability and construct and predictive validity;[20,30,31] the 11-item self-destructive behaviour supplement to the Borderline Symptom List (BSL;[32]), which assesses past week engagement in several self-destructive behaviours (e.g., binge eating, excessive drinking, drug-use, risky sexual behaviours); the 21-item Depression Anxiety Stress Scales (DASS-21;[33]), a measure of depression, anxiety, and stress symptoms (α=.88 in this sample) with good test-retest reliability and construct and discriminant validity;[33,34] the BPD-related composite of the Inventory of Interpersonal

Problems (IIP-BPD;[35]), a measure of BPD-relevant interpersonal difficulties (α=.89 in this sample) with good convergent validity and specificity;[35] and the Sheehan Disability Scales (SDS;[36]), a widely used measure of social and vocational impairment due to psychological symptoms (α=.80 in this sample) with adequate reliability and construct, convergent, and discriminant validity across various clinical populations.[37,38] Treatment credibility and expectancy were assessed after the second session of ERGT with the Credibility/Expectancy Questionnaire.[39]

Treatment

ERGT is a 14-session, adjunctive, acceptance-based behavioural group treatment developed to treat DSH by targeting its underlying mechanism of emotion dysregulation.[11] Based on Gratz and Roemer's[30] conceptualization of emotion regulation, ERGT systematically teaches skills aimed at improving a number of dimensions of emotion regulation, including: emotional awareness, understanding, and acceptance; the ability to control behaviours when experiencing negative emotions; the use of non-avoidant emotion regulation strategies to modulate the intensity and/or duration of emotional responses; and the willingness to experience negative emotions as part of pursuing meaningful activities in life. Moreover, the following themes are emphasized throughout the treatment: (a) the potentially paradoxical effects of emotional avoidance, (b) the emotion regulating consequences of emotional acceptance and willingness, and (c) the importance of controlling behaviour when emotions are present, rather than controlling emotions themselves. A detailed description of the content and development of ERGT is available elsewhere.[11]

The ERGT treatment manual (Gratz & Tull, 2010, unpublished manual) provides thorough instructions on the therapeutic stance and theoretical framework of ERGT, as well as detailed descriptions of the content, in-session exercises, and homework assignments for each session. Sudden deterioration or suicidal crises are monitored weekly through self-report measures assessing DSH frequency and emotion dysregulation, and addressed by the group therapists in collaboration with the ongoing treatment provider if needed. The ERGT treatment manual was translated into Swedish through a collaborative and iterative process involving the treatment developer (KLG) and co-author of the ERGT manual (MTT) and the primary ERGT supervisors for this trial (HS and JB, in collaboration with LGL). In this trial, ERGT was delivered in weekly two-hour sessions led by two therapists with groups of 4-9 participants (average group size = 5.59).

Therapist Training and Treatment Fidelity

Prior to the trial, therapists were required to study the ERGT manual and relevant articles on the theoretical underpinnings of ERGT, as well as to participate in a three-day workshop led by the authors of the ERGT manual (KLG and MTT). The workshop consisted of didactic lectures, demonstrations, role-playing, and practice exercises. Of the 28 therapists, 22 were licensed psychologists, two were social workers, two were nurses, one was a psychiatric aid, and one was an occupational therapist. Twenty-six therapists had previous experience treating patients with BPD and/or DSH; three therapists had previous experience with ACT, and 19 therapists had experience with DBT. To ensure treatment fidelity during the ERGT trial, all sessions were filmed and reviewed weekly by clinicians with experience delivering ERGT (HS, JB), and all therapists received the option of weekly supervision based on the reviewed films.

Statistics and Data Analysis

We expected to recruit 90-105 participants. This sample size yielded high power (> .99 with $\alpha = .05$) to detect a standardized mean difference between pre- and post-treatment of $d = 0.5$ using within-group t-tests on a log-transformation of the primary outcome measure DSH-frequency (consistent with past research on ERGT;[11,12,20]).

All analyses were performed in R using random effects modelling.[40] The count variables, DSH frequency and BSL, are reported with medians and interquartile range and were analysed using negative binomial generalized mixed models, and the remaining continuous outcomes are reported as means and standard deviations and were analysed using linear models. The models included all available data at the three assessment points (pre-treatment, post-treatment, and six-month follow-up) for each outcome, thus making them intent-to-treat analyses. We estimated separate slopes for the change between the pre- and post-treatment assessments (S1) and the change between the post-treatment and six-month follow-up assessments (S2). Random intercepts and random slopes as well as autoregressive correlation structures were included in the models if they significantly improved model fit according to log-likelihood ratio tests.

Effect sizes were calculated for changes between pre-treatment, post-treatment, and six-month follow-up. For the count variables (i.e. DSHI-frequency and BSL), the percentage change from baseline to any subsequent time point with 95% confidence intervals was used as an effect size. This was calculated by exponentiating the estimate for the slopes derived from the negative binomial models, and interpreting the range below or above one as the percentage decrease or

increase in the outcome for a one-unit increase in the predictor. Effect sizes for the remaining continuous outcomes are reported as Cohen's *d,* calculated by dividing the appropriate slope estimate (i.e., pre- to post-treatment: S1, pre-treatment to six-month follow-up: S1+S2, and post-treatment to six-month follow-up: S2) by the pre-treatment standard deviation. We also performed separate analyses where linear mixed models were applied to log-transformed DSH frequency and BSL scores and corresponding effect sizes were extracted. These effect sizes are reported together with observed means and standard deviations of DSH frequency and BSL scores to permit comparison with previous studies of ERGT.[11,12,20] However, inferences of the statistical significance of changes on these measures were based on the more appropriate negative binomial regression models. Confidence intervals with a 95% margin for the effect sizes were calculated using 5000 bootstrap replications.[41] The bootstrap replications were clustered on participants.[42]

We also performed sensitivity analyses to investigate the robustness and validity of the DSHI results. First, the treating clinics were entered as random factors to test the possibility of clustering effects in the data. Second, we added concurrent medication status (coded as 0 for no concurrent medication and 1 for concurrent medication) and type of treatment as usual (coded as a factor with the following levels: cognitive behavioural therapy, psychodynamic therapy, supportive therapy, or other) as covariates in the model. These covariates were added both as simple effects and as interaction effects with the S1 variable to investigate the possible influence of the covariates on DSH-frequency and treatment effect. Third, we entered the number of treatment sessions attended as a predictor of improvement in DSH frequency during the treatment period.

Finally, we examined the number of participants who reported no (zero) DSH episodes at each assessment point and used McNemar's exact tests to analyse the changes between the assessment points.

RESULTS

Treatment Adherence and Attrition

The average time from baseline assessment to the start of treatment was 20.7 days (SD = 17.2, range 1-91). Twenty-one participants (22%) dropped out of ERGT (see Figure 1). Mean number of sessions attended for all included participants was 11 (SD = 5.2, min 0, max 16). Seventy-two participants (76%) attended ≥ 7 sessions and 47 (49%) attended 14 sessions. Post-treatment assessments were completed by 88 (93%) participants and six-month follow-up assessments were completed by 76 (82%) participants. Mean ratings of treatment credibility and expectancy completed after the second session were 5.7 (SD = 1.8) and 47.0% (SD = 25.5), respectively (for comparison, mean ratings reported by Gratz & Tull were 6.91 and 57%, respectively).[20]

Primary Outcome

Results for both the primary and secondary outcome measures are displayed in Table 2. There was a significant 52% reduction in DSH frequency from pre- to post-treatment, and a 76% reduction from pre-treatment to six-month follow-up.

The observed means for DSH frequency were 53.68 (SD = 99.88), 37.45 (SD = 72.22), and 28.69 (SD = 95.44) at pre-treatment, post-treatment, and follow-up, respectively. Effect sizes

(*Cohen's d*) based on log-transformed data showed medium-sized reductions in DSH frequency from pre- to post-treatment (d = 0.52, 95% CI: 0.30, 0.75) and from post-treatment to six-month follow-up (d = 0.47, 95% CI: 0.27, 0.70). A large effect size was observed from pre-treatment to six-month follow-up (d = 0.99, 95% CI: 0.70, 1.30).

The percentage of participants with an observed score of zero DSH episodes during the past four months increased significantly from 4.2% at pre-treatment to 17.9% at post-treatment (Exact McNemar's χ^2 = 12.25, df = 1, $p < .001$) and to 25.3% at follow-up (Exact McNemar's χ^2 = 16.67, df = 1, $p < .001$). The increase between post-treatment and follow-up was also significant (Exact McNemar's χ^2 = 7.14, df = 1, $p = .01$).

Secondary Outcomes

There were significant improvements in DSH versatility, emotion dysregulation, self-destructive behaviours, and depression and stress symptoms at post-treatment, accompanied by small to large effect sizes (see Table 2). At the six-month follow-up, all of these improvements were either maintained or further improved upon, with the change on the DERS from pre-treatment to six-month follow-up reaching a large-sized effect. Median reduction in self-destructive behaviours over each time period is reported in Table 2. The observed mean self-destructive behaviour scores were 4.82 (SD = 3.69), 3.65 (SD = 4.24) and 3.24 (SD = 3.61) at pre-treatment, post-treatment, and follow-up, respectively. Analyses on log-transformed data showed small to medium effect sizes between pre- and post-treatment (*Cohen's d* = 0.43, 95% CI: 0.20, 0.65) and pre- and six-month follow-up (d = 0.55, 95% CI: 0.27, 0.77). Improvements in interpersonal difficulties, anxiety, and social and vocational impairment were not significant at post-treatment;

however, at six-month follow-up, there were significant improvements in these outcomes (accompanied by small to moderate effect sizes) from pre-treatment (stemming from the additional significant improvements in these measures observed from post-treatment to six-month follow-up).

Sensitivity analyses

First, we included the treating clinics as random factors in the DSH frequency model. This did not improve model fit, suggesting that there was not a significant clustering effect of treating clinic in the data. Second, we included concurrent medication status and type of treatment as usual as simple effects and interaction effects with S1 (i.e., the change between the pre-treatment and post-treatment assessments) in the DSH frequency model. None of the added predictors were statistically significant ($p > .05$), suggesting that concurrent medication use and type of treatment as usual were not associated with treatment effect. Third, we included the number of attended sessions as a predictor of improvement in DSH frequency. The interaction effect between session count and S1 was statistically significant, ($B = 0.06$, $Z = 11.54$, $p < .001$), indicating that higher attendance was associated with larger improvements during treatment.

Table 2. Treatment outcome variables for count- and continuous outcomes at pre-treatment, post-treatment and six-month follow-up

Outcome *Count-data*	Pre-treatment	Post-treatment	6-mo f-u	Pre-to post-treatment comparison		Post- to 6-mo follow-up comparison		Pre-to 6-mo follow-up comparison		Model *p*[b]
	Median (IQR)	*Median (IQR*	*Median (IQR*	*Z*	*Percent change*[a] *[95% CI]*	*Z*	*Percent change*[a] *[95% CI]*	*Z*	*Percent change*[a] *[95% CI]*	
DSHI-f	22.00 (9.5-56.0)	10.00 (2.8-45.5)	4.0 (0.0-13.0)	0.24***	52% [0.33, 0.66]	3.62***	49% [27, 65]	7.33***	76% [65, 83]	<.001
BSL	4.00 (2.0-7.0)	2.0 (1.0-5.0)	2.0 (0.3-5.0)	3.60***	29% [14, 41]	0.76	8% [-14, 25]	4.18***	34% [20. 46]	<.001
Outcome *Continuous*	Pre-treatment	Post-treatment	6-mo f-u	Pre-to post-treatment comparison		Post- to 6-mo follow-up comparison		Pre-to 6-mo follow-up comparison		Model *p*
	Mean (SD)	*Mean (SD)*	*Mean (SD)*	*Z*	*Cohen's d*[a] *[95% CI]*	*Z*	*Cohen's d*[a] *[95% CI]*	*Z*	*Cohen's d*[a] *[95% CI]*	
DSHI-v	3.01 (1.82)	2.23 (2.02)	1.67 (1.76)	3.67***	0.41 [0.19, 0.63]	2.86**	0.24 [0.07, 0.42]	5.65***	0.65 [0.40, 0.89]	<.001
DERS	125.98 (19.37)	108.17 (27.52)	104.66 (27.40)	6.56***	0.91 [0.63, 1.20]	1.00	0.12 [-0.09, 0.43]	7.13***	1.03 [0.69, 1.38]	<.001
IIP-BPD	2.05 (0.72)	1.98 (0.71)	1.78 (0.76)	1.08	0.10 [-0.06, 0.29]	2.43*	0.24 [0.06, 0.43]	3.46***	0.34 [0.13, 0.59]	.002
DASS-D	25.35 (10.28)	20.11 (11.80)	19.95 (11.92)	4.37***	0.50 [0.29, 0.75]	0.41	0.05 [-0.22, 0.30]	3.91***	0.55 [0.27, 0.85]	<.001
DASS-A	17.14 (8.98)	16.30 (9.97)	14.54 (9.58)	0.85	0.08 [-0.10, 0.27]	1.59	0.17 [-0.05, 0.41]	2.11*	0.25 [0.01, 0.49]	.024
DASS-S	25.77 (7.95)	23.34 (9.21)	21.19 (10.34)	2.46*	0.30 [0.06, 0.54]	1.95*	0.26 [0.01, 0.52]	4.30***	0.56 [0.26, 0.86]	<.001
SDS	18.44 (6.99)	18.59 (7.05)	16.01 (8.41)	-0.25	-0.03 [-0.22, 0.16]	2.71**	0.32 [0.07, 0.60]	2.51*	0.29 [0.04, 0.59]	.013

Note. Test statistics are based on negative binomial generalized mixed models for count data and linear mixed models analyses for continuous data. Confidence intervals for effect sizes are based on 5000 bootstrap replications. Abbreviations: BSL = Borderline Symptom List, behaviour supplement, DASS-D = Depression Anxiety and Stress Scales – 21 Depression, DASS-A = Depression Anxiety and Stress Scales – 21 Anxiety, DASS-S = Depression Anxiety and Stress Scales – 21 Stress, DERS = Difficulties in Emotion Regulation Scale, DSHI-f = Deliberate Self-Harm Inventory - frequency, DSHI-v = Deliberate Self-Harm Inventory - versatility, IIP-BPD = BPD-related composite of the Inventory of Interpersonal Problems, SDS = Sheehan Disability Scales.

[a]Effect size estimates for DSHI frequency and BSL are reported as percent change across time, and effect sizes for the remaining continuous outcomes are reported as *Cohen's d*.

[b]*p*-value of log-likelihood ratio comparison with null model including no fixed effects and only random intercept.

*p<.05, **p<.01, ***p<.001

DISCUSSION

The present multi-site evaluation of ERGT in routine clinical care provides additional support for the feasibility and transportability of this treatment. Results revealed significant improvements in DSH frequency and versatility, emotion dysregulation, self-destructive behaviours, and depression and stress symptoms from pre- to post-treatment. By the six-month follow-up, interpersonal difficulties and social and vocational impairment had also improved significantly. Moreover, all gains found at post-treatment were either maintained or further improved upon at follow-up. Notably, the results of the sensitivity analyses strengthen our confidence in these findings, providing evidence of a significant effect of session attendance, but not treatment clinic, concurrent medication use, or type of treatment as usual, on improvements in DSH frequency. These results are consistent with past findings that characteristics of participants' ongoing therapy in the community had minimal impact on treatment response to ERGT,[43] and suggest that it is engagement in ERGT rather than other (non-specific) treatment-related factors that influences reductions in DSH frequency. Nonetheless, the uncontrolled open trial design precludes conclusions regarding the causal relation of treatment participation to symptom improvement as there may be other factors underlying both treatment completion status and symptom improvement (i.e., motivation for treatment or alliance with treatment providers) that may, at least in part, account for the present findings.

Although the results of this study need to be interpreted with caution (due to the lack of a control group), they are similar to those obtained in previous ERGT efficacy trials,[11,12]. Likewise, our

findings of either stability or further improvements during the follow-up period are consistent with the observed pattern of continued or maintained improvement during a 9-month follow-up period in Gratz and colleagues' RCT.[12] These findings provide further support for the durability of improvements following this relatively brief and non-intensive treatment, suggesting that participants may experience durable gains through ERGT even when delivered by community clinicians with only brief training in this treatment.

Despite these similarities with previous ERGT trials, the pattern of findings for measures of interpersonal, social, and vocational functioning differed from Gratz and colleagues' previous ERGT studies.[12,20] Specifically, results revealed no significant changes in interpersonal difficulties or social and vocational impairment from pre- to post-treatment, although there were significant (albeit small) improvements in these areas during the follow-up period. Although the differences in these findings may be due to study-related differences in therapist training and/or treatment delivery, they may also be explained (in part) by the fact that more than one-third of participants (39%) in the present trial were on disability pension or current sick leave, thus limiting their social interactions and vocational opportunities during treatment.

Notably, rates of abstinence from DSH increased significantly from pre- to post-treatment, as well as from post-treatment through the six-month follow-up, with 25.2% of participants reporting abstinence from DSH six-months post-treatment. Nonetheless, it warrants mention that the proportion of participants reaching abstinence from DSH was lower than in previous ERGT trials.[12,20] This may be due to differences in the level of training and supervision provided in this trial versus previous trials. Indeed, ratings of treatment credibility and expectancy in this trial

were somewhat lower than in previous studies of ERGT, potentially capturing the lesser experience of the therapists in this trial. Conversely, the lower rate of treatment-related abstinence from DSH observed in this study may reflect differences in sample composition and/or clinical severity (as both emotion dysregulation and DSH frequency reported in this sample were at the high end of the range reported in past ERGT studies). Nonetheless, past research examining predictors of treatment response to ERGT found that several indicators of greater severity in domains relevant to ERGT (i.e., baseline emotion dysregulation and BPD criteria, lifetime and recent DSH, and past-year hospitalization and suicide attempts) predicted *better* responses during treatment and follow-up,[43] suggesting that greater severity in certain domains may be associated with better response to this treatment. Future research is needed to clarify the particular patients most likely to benefit from ERGT.

Importantly, we found significant reductions across all assessment points in DSH versatility (i.e., a marker of DSH severity associated with future suicide risk).[29] This finding lends support to the potential utility of ERGT in decreasing risk for self-injurious behaviours in general. Indeed, given the high occurrence of DSH and suicidal behaviours,[9] as well as evidence that DSH is one of the strongest prospective predictors of future suicide attempts,[44,45] the emphasis within ERGT on reducing DSH could be expected to reduce suicidal risk as well. Nonetheless, in the absence of data in this or previous ERGT trials on suicidal outcomes in particular, the benefits of ERGT for suicidality remain unknown and in need of future investigation.

Although results of this study provide preliminary support for the utility of ERGT in routine clinical care, they also highlight more broadly the potential utility of targeting emotion

dysregulation in the treatment of BPD-related pathology. Notably, there are several other empirically-supported treatments for BPD and related pathology that directly target emotion dysregulation, including DBT[6] and systems training for emotional predictability and problem solving (STEPPS;[46]). Findings of the efficacy of those treatments for BPD,[24,46,47] combined with both the results of the current study and past findings supporting the efficacy of ERGT for DSH within BPD,[11,12] highlight the potential benefits of interventions aimed at promoting adaptive emotion regulation among individuals with BPD and provide further support for emotion dysregulation as a primary mechanism underlying the pathogenesis and effective treatment of BPD and related pathology.[48]

There are several strengths of this study that are worth noting. First, the naturalistic design of this study permitted inclusion of a large sample of participants who received this ERGT as part of their standard treatment. Second, data attrition was low (8% at post-treatment and 18% at follow-up) and participant dropout was within expected levels when treating individuals with BPD[49] and consistent with previous ERGT trials.[12] Third, contrary to previous ERGT trials,[11,12,20] the group leaders were of different professional backgrounds and representative of the community clinicians who regularly treat this patient population, providing support for the generalizability and transportability of this treatment. Finally, our data provide support for the feasibility of disseminating ERGT to community clinicians, as therapists in this study were provided with only readings and a brief workshop on ERGT prior to its delivery.

Despite these strengths, findings must be considered in light of limitations present. First, the lack of a control condition precludes conclusions about the effects of this ERGT (vs. treatment as

usual or the passage of time). However, it is important to note that the waitlist conditions in previous ERGT trials generally evidenced stability over time on all measures of interest (likely due to the relatively short time-frame of the treatment period). Nonetheless, future studies are needed to evaluate the effectiveness of ERGT within an RCT design. Second, even though the participating clinicians were offered weekly supervision based on filmed sessions, no systematic adherence ratings were conducted. This limits our ability to speak to the quality of ERGT and its relation to outcome. Third, the results relied solely on self-report measures and not clinician-rated outcomes. However, research on self-reported versus clinician-rated outcomes in psychotherapy studies suggests that the use of self-report measures often results in smaller effect sizes when compared to clinician administered instruments.[50] Thus, it is possible that our results may be conservative estimates of improvements during and after treatment. Fourth, our study only included women, which limits the generalizability of the results to men. Future research is needed that evaluates ERGT within male or mixed-gender samples. Finally, current (past month) substance dependence was an exclusion criterion in this study, both to ensure comparability with previous ERGT trials and because active substance dependence requires specialized treatment and a higher level of care. Indeed, in Sweden, individuals with ongoing substance dependence are not necessarily offered treatment within a psychiatric clinic; rather, they may be referred to specialized substance dependence treatment centers. Thus, it is unlikely that the exclusion of such individuals negatively affected the generalizability of this sample to a typical Swedish psychiatric clinic. Furthermore, this criterion did not exclude individuals with past (or even recent) substance use problems. Nonetheless, given the high co-occurrence of substance dependence and BPD,[51] excluding these individuals may have limited the external validity of the sample to patients with BPD as a whole.

Despite these limitations, our results provide further evidence for the utility and transportability of ERGT, suggesting that this is a feasible treatment for DSH, emotion dysregulation, and psychiatric symptoms when delivered by community clinicians in traditional clinical settings.

ACKNOWLEDGEMENTS

The authors wish to thank Ida Janson, Sara Eivergård, Niclas Andersson and all study therapists for their hard work and enthusiasm, which made this study possible.

COMPETING INTERESTS

Ms Hanna Sahlin, Mr Johan Bjureberg, Dr Kim L Gratz, Dr Matthew T Tull, Dr Jonas Bjärehed, Dr Lars-Gunnar Lundh, Dr Jussi Jokinen and Dr Clara Hellner Gumpert report no competing interests. Dr Erik Hedman and Dr Brjánn Ljótsson are shareholders of a company, Dahlia, specialized in psychiatric symptom assessment.

FUNDING

This research was supported by the National Self Harm project in Sweden and Stockholm County Council regional research grant #SLL20140428. None of the funding organizations has had any role in the design and conduct of the study; in the collection, management, and analysis of the data; or in the preparation, review and approval of the manuscript.

DATA SHARING

No additional data are available.

AUTHOR CONTRIBUTIONS

Drs Sahlin, Bjureberg, Ljótsson and Gumpert had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Drs Sahlin, Bjureberg, Gratz, Tull, Hedman, Bjärehed, Lundh, Jokinen, Ljótsson and Gumpert. Drs Sahlin, Bjureberg, Bjärehed, Ljótsson and Gumpert

Drafting of the manuscript: Drs Sahlin, Bjureberg, Ljótsson and Gumpert

Critical revision of the manuscript for important intellectual content and approval of the final version to be published: Drs Sahlin, Bjureberg, Gratz, Tull, Hedman, Bjärehed, Lundh, Jokinen, Ljótsson and Gumpert

TRIAL REGISTRATION

This trial was registered on Clinicaltrials.gov on November 4, 2013.

Clinicaltrials.gov identifier NCT01986257

REFERENCES

1 Gratz KL. Measurement of Deliberate Self-Harm: Preliminary Data on the Deliberate Self-Harm Inventory. *J Psychopathol Behav Assess* 2001;**23**:253–63. doi:10.1023/A%3A1012779403943

2 Briere J, Gil E. Self-mutilation in clinical and general population samples: prevalence, correlates, and functions. *Am J Orthopsychiatry* 1998;**68**:609–20.

3 de Klerk S, van Noorden MS, van Giezen AE, *et al.* Prevalence and correlates of lifetime deliberate self-harm and suicidal ideation in naturalistic outpatients: The Leiden Routine Outcome Monitoring study. *J Affect Disord* 2011;**133**:257–64. doi:10.1016/j.jad.2011.03.021

4 Swannell SV, Martin GE, Page A, *et al.* Prevalence of nonsuicidal self-injury in nonclinical samples: systematic review, meta-analysis and meta-regression. *Suicide Life Threat Behav* 2014;**44**:273–303. doi:10.1111/sltb.12070

5 Sansone RA, Wiederman MW, Sansone LA. The Self-Harm Inventory (SHI): development of a scale for identifying self-destructive behaviors and borderline personality disorder. *J Clin Psychol* 1998;**54**:973–83.

6 Linehan M. *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford Press 1993.

7 Zanarini MC. Psychotherapy of borderline personality disorder. *Acta Psychiatr Scand* 2009;**120**:373–7. doi:10.1111/j.1600-0447.2009.01448.x

8 Victor SE, Klonsky ED. Correlates of suicide attempts among self-injurers: A meta-analysis. *Clin Psychol Rev* 2014;**34**:282–97. doi:10.1016/j.cpr.2014.03.005

9 Hamza CA, Stewart SL, Willoughby T. Examining the link between nonsuicidal self-injury and suicidal behavior: a review of the literature and an integrated model. *Clin Psychol Rev* 2012;**32**:482–95. doi:10.1016/j.cpr.2012.05.003

10 Bateman AW, Fonagy P. Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *Am J Psychiatry* 1999;**156**:1563–9. doi:10.1176/ajp.156.10.1563

11 Gratz KL, Gunderson JG. Preliminary data on an acceptance-based emotion regulation group intervention for deliberate self-harm among women with borderline personality disorder. *Behav Ther* 2006;**37**:25–35. doi:10.1016/j.beth.2005.03.002

12 Gratz KL, Tull MT, Levy R. Randomized 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. *Psychol Med* 2014;**44**:2099–112. doi:10.1017/S0033291713002134

13 Bales D, van Beek N, Smits M, *et al.* Treatment outcome of 18-month, day hospital mentalization-based treatment (MBT) in patients with severe borderline personality disorder in the Netherlands. *J Pers Disord* 2012;**26**:568–82. doi:10.1521/pedi.2012.26.4.568

14 Feigenbaum JD. A real-world study of the effectiveness of DBT in the UK National Health Service. *Br J Clin Psychol* 2012;**51**:121–41. doi:10.1111/j.2044-8260.2011.02017.x

15 Pasieczny N, Connor J. The effectiveness of dialectical behaviour therapy in routine public mental health settings: an australian controlled trial. *Behav Res Ther* 2011;**49**:4–10. doi:10.1016/j.brat.2010.09.006

16 Wagner T, Fydrich T, Stiglmayr C, *et al.* Societal cost-of-illness in patients with borderline personality disorder one year before, during and after dialectical behavior therapy in routine outpatient care. *Behav Res Ther* 2014;**61**:12–22. doi:10.1016/j.brat.2014.07.004

17 Swales MA, Taylor B, Hibbs RAB. Implementing Dialectical Behaviour Therapy: programme survival in routine healthcare settings. *J Ment Health* 2012;**21**:548–55. doi:10.3109/09638237.2012.689435

18 Weisz JR, Weiss B, Donenberg GR. The lab versus the clinic. Effects of child and adolescent psychotherapy. *Am Psychol* 1992;**47**:1578–85.

19 Wells KB. Treatment Research at the Crossroads: The Scientific Interface of Clinical Trials and Effectiveness Research. *Am J Psychiatry* 1999;**156**:5–10.

20 Gratz KL, Tull MT. Extending research on the utility of an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality pathology. *Personal Disord* 2011;**2**:316–26. doi:10.1037/a0022144

21 First MB, Gibbon M, Spitzer RL, *et al. Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II)*. Washington, D.C: American Psychiatric Press, Inc. 1997.

22 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (4th ed., Text Revision).* Washinghon, DC: Author 2000.

23 Sheehan DV, Lecrubier Y, Sheehan KH, *et al.* The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;**59 Suppl 20**:22–33–quiz34–57.

24 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**:757–66. doi:10.1001/archpsyc.63.7.757

25 McMain SF, Links PS, Gnam WH, *et al.* A Randomized Trial of Dialectical Behavior Therapy Versus General Psychiatric Management for Borderline Personality Disorder. *Am J Psychiatry* 2009;**166**:1365–74. doi:10.1176/appi.ajp.2009.09010039

26 Statistics Sweden. Population in the Country, Counties and Municipalities on 31/12/2013 and Population Change in 2013. 2014. http://www.scb.se/en_/Yearly-statistics--Municipalities-Counties-and-the-whole-country/370301/

27 Posner K, Brown GK, Stanley BH, *et al.* The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry* 2011;**168**:1266–77. doi:10.1176/appi.ajp.2011.10111704

28 Hedman E, Ljótsson B, Andersson E, *et al.* Effectiveness and cost offset analysis of group CBT for hypochondriasis delivered in a psychiatric setting: an open trial. *Cogn Behav Ther* 2010;**39**:239–50. doi:10.1080/16506073.2010.496460

29 Turner BJ, Layden BK, Butler SM, *et al.* How often, or how many ways: clarifying the relationship between non-suicidal self-injury and suicidality. *Arch Suicide Res* 2013;**17**:397–415. doi:10.1080/13811118.2013.802660

30 Gratz KL, Roemer L. Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure, and Initial Validation of the Difficulties in Emotion Regulation Scale. *J Psychopathol Behav Assess* 2004;**26**:41–54. doi:10.1023/B%3AJOBA.0000007455.08539.94

31 Bjureberg J, Ljótsson B, Tull MT, *et al.* Development and Validation of a Brief Version of the Difficulties in Emotion Regulation Scale: The DERS-16. *J Psychopathol Behav Assess* 2015;:1–13. doi:10.1007/s10862-015-9514-x

32 Bohus M, Limberger MF, Frank U, *et al.* [Development of the Borderline Symptom List]. *Psychother Psychosom Med Psychol* 2001;**51**:201–11. doi:10.1055/s-2001-13281

33 Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behav Res Ther* 1995;**33**:335–43.

34 Antony MM, Bieling PJ, Cox BJ, *et al.* Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychol Assess* 1998;**10**:176–81. doi:10.1037/1040-3590.10.2.176

35 Lejuez CW, Daughters SB, Nowak JA, *et al.* Examining the inventory of interpersonal problems as a tool for conducting analogue studies of mechanisms underlying Borderline Personality Disorder. *J Behav Ther Exp Psychiatry* 2003;**34**:313–24. doi:10.1016/j.jbtep.2003.11.002

36 Sheehan DV. *The Sheehan disability scales. The anxiety disease and how to overcome it.* New York City: Charles Scribner and Sons 1983.

37 Hambrick JP, Turk CL, Heimberg RG, *et al.* Psychometric properties of disability measures among patients with social anxiety disorder. *J Anxiety Disord* 2004;**18**:825–39. doi:10.1016/j.janxdis.2003.10.004

38 Diefenbach GJ, Abramowitz JS, Norberg MM, *et al.* Changes in quality of life following cognitive-behavioral therapy for obsessive-compulsive disorder. *Behav Res Ther* 2007;**45**:3060–8. doi:10.1016/j.brat.2007.04.014

39 Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry* 2000;**31**:73–86.

40 R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: 2015. http://www. R-project.org/

41 Kelley K. The Effects of Nonnormal Distributions on Confidence Intervals Around the Standardized Mean Difference: Bootstrap and Parametric Confidence Intervals. *Edu Psychol Meas* 2005;**65**:51–69. doi:10.1177/0013164404264850

42 Ren S, Lai H, Tong W, *et al.* Nonparametric bootstrapping for hierarchical data. *J Appl Stat* 2010;**37**:1487–98. doi:10.1080/02664760903046102

43 Gratz KL, Dixon-Gordon KL, Tull MT. Predictors of treatment response to an adjunctive emotion regulation group therapy for deliberate self-harm among women with borderline personality disorder. *Personal Disord* 2014;**5**:97–107. doi:10.1037/per0000062

44 Asarnow JR, Porta G, Spirito A, *et al.* Suicide attempts and nonsuicidal self-injury in the treatment of resistant depression in adolescents: findings from the TORDIA study. *J Am Acad Child Adolesc Psychiatry* 2011;**50**:772–81. doi:10.1016/j.jaac.2011.04.003

45 Wilkinson P, Kelvin R, Roberts C, *et al.* Clinical and psychosocial predictors of suicide attempts and nonsuicidal self-injury in the Adolescent Depression Antidepressants and Psychotherapy Trial (ADAPT). *Am J Psychiatry* 2011;**168**:495–501. doi:10.1176/appi.ajp.2010.10050718

46 Blum N, St John D, Pfohl B, *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**:468–78. doi:10.1176/appi.ajp.2007.07071079

47 Linehan MM, Korslund KE, Harned MS, *et al.* Dialectical Behavior Therapy for High Suicide Risk in Individuals With Borderline Personality Disorder. *JAMA Psychiatry* 2015;**72**:475–8. doi:10.1001/jamapsychiatry.2014.3039

48 Gratz KL, Weiss NH, Tull MT. Examining Emotion Regulation as an Outcome, Mechanism, or Target of Psychological Treatments. *Curr Opin Psychol* 2015;**3**:85–90. doi:10.1016/j.copsyc.2015.02.010

49 Barnicot K, Katsakou C, Marougka S, *et al.* Treatment completion in psychotherapy for borderline personality disorder: a systematic review and meta-analysis. *Acta Psychiatrica Scandinavica* 2011;**123**:327–38. doi:10.1111/j.1600-0447.2010.01652.x

50 Cuijpers P, Li J, Hofmann SG, *et al.* Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: a meta-analysis. *Clin Psychol Rev* 2010;**30**:768–78. doi:10.1016/j.cpr.2010.06.001

51 Trull TJ, Sher KJ, Minks-Brown C, *et al.* Borderline personality disorder and substance use disorders: a review and integration. *Clin Psychol Rev* 2000;**20**:235–53.

Figure 1. Participant flow through the study

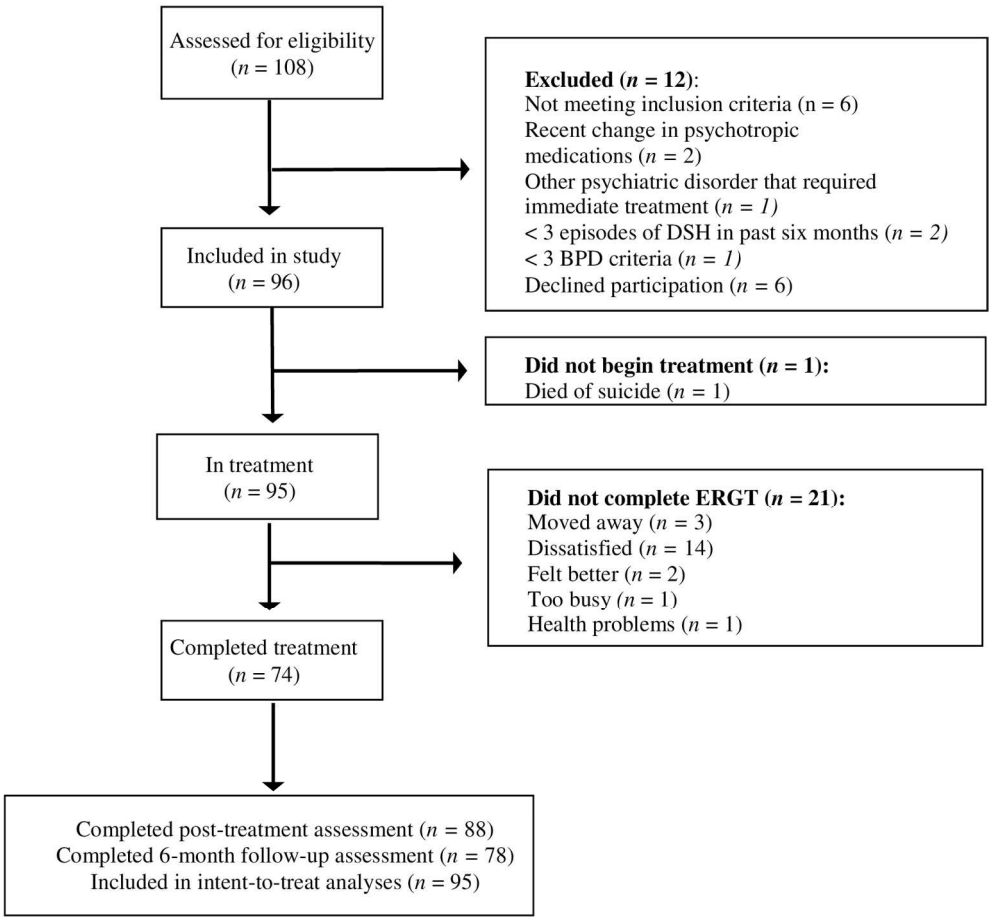

Figure 1. Participant flow through the study

157x162mm (300 x 300 DPI)

TREND Statement Checklist

Paper Section/ Topic	Item No	Descriptor	Reported? ✔	Reported? Pg #
Title and Abstract				
Title and Abstract	1	• Information on how unit were allocated to interventions	✔	2
		• Structured abstract recommended	✔	2
		• Information on target population or study sample	✔	2
Introduction				
Background	2	• Scientific background and explanation of rationale	✔	4
		• Theories used in designing behavioral interventions	✔	4, 9
Methods				
Participants	3	• Eligibility criteria for participants, including criteria at different levels in recruitment/sampling plan (e.g., cities, clinics, subjects)	✔	5
		• Method of recruitment (e.g., referral, self-selection), including the sampling method if a systematic sampling plan was implemented	✔	4
		• Recruitment setting	✔	4-5
		• Settings and locations where the data were collected	✔	6
Interventions	4	• Details of the interventions intended for each study condition and how and when they were actually administered, specifically including:		
		○ Content: what was given?	✔	8
		○ Delivery method: how was the content given?	✔	8
		○ Unit of delivery: how were the subjects grouped during delivery?	✔	7
		○ Deliverer: who delivered the intervention?	✔	7, 9
		○ Setting: where was the intervention delivered?	✔	6
		○ Exposure quantity and duration: how many sessions or episodes or events were intended to be delivered? How long were they intended to last?	✔	8
		○ Time span: how long was it intended to take to deliver the intervention to each unit?	✔	5
		○ Activities to increase compliance or adherence (e.g., incentives)		
Objectives	5	• Specific objectives and hypotheses	✔	5
Outcomes	6	• Clearly defined primary and secondary outcome measures	✔	8
		• Methods used to collect data and any methods used to enhance the quality of measurements	✔	8
		• Information on validated instruments such as psychometric and biometric properties	✔	9
Sample Size	7	• How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules	✔	10
Assignment Method	8	• Unit of assignment (the unit being assigned to study condition, e.g., individual, group, community)	✔	7, 9
		• Method used to assign units to study conditions, including details of any restriction (e.g., blocking, stratification, minimization)		
		• Inclusion of aspects employed to help minimize potential bias induced due to non-randomization (e.g., matching)		

TREND Statement Checklist

Blinding (masking)	9	• Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to study condition assignment; if so, statement regarding how the blinding was accomplished and how it was assessed.		
Unit of Analysis	10	• Description of the smallest unit that is being analyzed to assess intervention effects (e.g., individual, group, or community)	✔	11
		• If the unit of analysis differs from the unit of assignment, the analytical method used to account for this (e.g., adjusting the standard error estimates by the design effect or using multilevel analysis)	✔	10
Statistical Methods	11	• Statistical methods used to compare study groups for primary methods outcome(s), including complex methods of correlated data	✔	10
		• Statistical methods used for additional analyses, such as a subgroup analyses and adjusted analysis	✔	10
		• Methods for imputing missing data, if used	✔	10
		• Statistical software or programs used	✔	10
Results				
Participant flow	12	• Flow of participants through each stage of the study: enrollment, assignment, allocation, and intervention exposure, follow-up, analysis (a diagram is strongly recommended)	✔	21, Figure 1
		○ Enrollment: the numbers of participants screened for eligibility, found to be eligible or not eligible, declined to be enrolled, and enrolled in the study	✔	21
		○ Assignment: the numbers of participants assigned to a study condition	✔	21
		○ Allocation and intervention exposure: the number of participants assigned to each study condition and the number of participants who received each intervention	✔	21
		○ Follow-up: the number of participants who completed the follow-up or did not complete the follow-up (i.e., lost to follow-up), by study condition	✔	21, 11
		○ Analysis: the number of participants included in or excluded from the main analysis, by study condition	✔	21, 11
		• Description of protocol deviations from study as planned, along with reasons		
Recruitment	13	• Dates defining the periods of recruitment and follow-up	✔	5
Baseline Data	14	• Baseline demographic and clinical characteristics of participants in each study condition	✔	6
		• Baseline characteristics for each study condition relevant to specific disease prevention research		
		• Baseline comparisons of those lost to follow-up and those retained, overall and by study condition	✔	10
		• Comparison between study population at baseline and target population of interest	✔	16-17
Baseline equivalence	15	• Data on study group equivalence at baseline and statistical methods used to control for baseline differences		

TREND Statement Checklist

Numbers analyzed	16	• Number of participants (denominator) included in each analysis for each study condition, particularly when the denominators change for different outcomes; statement of the results in absolute numbers when feasible	✔	
		• Indication of whether the analysis strategy was "intention to treat" or, if not, description of how non-compliers were treated in the analyses	✔	10
Outcomes and estimation	17	• For each primary and secondary outcome, a summary of results for each estimation study condition, and the estimated effect size and a confidence interval to indicate the precision	✔	14
		• Inclusion of null and negative findings	✔	13
		• Inclusion of results from testing pre-specified causal pathways through which the intervention was intended to operate, if any		
Ancillary analyses	18	• Summary of other analyses performed, including subgroup or restricted analyses, indicating which are pre-specified or exploratory	✔	11
Adverse events	19	• Summary of all important adverse events or unintended effects in each study condition (including summary measures, effect size estimates, and confidence intervals)		
DISCUSSION				
Interpretation	20	• Interpretation of the results, taking into account study hypotheses, sources of potential bias, imprecision of measures, multiplicative analyses, and other limitations or weaknesses of the study	✔	16
		• Discussion of results taking into account the mechanism by which the intervention was intended to work (causal pathways) or alternative mechanisms or explanations		
		• Discussion of the success of and barriers to implementing the intervention, fidelity of implementation	✔	16,17
		• Discussion of research, programmatic, or policy implications	✔	17-18
Generalizability	21	• Generalizability (external validity) of the trial findings, taking into account the study population, the characteristics of the intervention, length of follow-up, incentives, compliance rates, specific sites/settings involved in the study, and other contextual issues	✔	17-18
Overall Evidence	22	• General interpretation of the results in the context of current evidence and current theory	✔	17-18

From: Des Jarlais, D. C., Lyles, C., Crepaz, N., & the Trend Group (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health*, 94, 361-366. For more information, visit: http://www.cdc.gov/trendstatement/

Correction: Emotion regulation group therapy for deliberate self-harm: a multi-site evaluation in routine care using an uncontrolled open trial design

Sahlin H, Bjureberg J, Gratz KL, *et al*. Emotion regulation group therapy for deliberate self-harm: a multi-site evaluation in routine care using an uncontrolled open trial design. *BMJ Open* 2017;**7**:e016220. doi: 10.1136/bmjopen-2017-016220.

Authors Hanna Sahlin and Johan Bjureberg contributed equally to the manuscript.

BMJ Open 2017;**7**:e016220corr1. doi:10.1136/bmjopen-2017-016220corr1