ACT-Enhanced Behavior Therapy in Group Format for Trichotillomania: An Effectiveness Study

Ashild Tellefsen Haaland  
Sørlandet Hospital

Shirin O. Eskeland  
Oslo University Hospital

Erna M. Moen  
Oslo University Hospital

Patrick A. Vogel  
Norwegian University of Science and Technology

Svein Haseth  
Østmarka Hospital

See next page for additional authors

Authors
Ashild Tellefsen Haaland, Shirin O. Eskeland, Erna M. Moen, Patrick A. Vogel, Svein Haseth, Kjetil Mellingen, Joseph A. Himle, Douglas W. Woods, and Benjamin Hummelen

This article is available at e-Publications@Marquette: https://epublications.marquette.edu/psych_fac/287
ACT-Enhanced Behavior Therapy in Group Format for Trichotillomania: An Effectiveness Study

Åshild Tellefsen Haaland
Department of Child and Adolescent Mental Health (ABUP), Sørlandet Hospital HF, Kristiansand, Norway

Shirin O. Eskeland
Clinic of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

Erna M. Moen
Clinic of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

Patrick A. Vogel
Norwegian University of Science and Technology, Psykologisk institutt, Trondheim, Norway

Svein Haseth
Østmarka hospital, Trondheim, Norway

Kjetil Mellingen
Clinic of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

Joseph A. Himle
School of Social Work and School of Medicine-Psychiatry, University of Michigan, Ann Arbor, MI
Highlights
• The results showed significant decreases in TTM symptoms from pre-treatment to the one-year follow-up.
• Nearly 90% of TTM patients scored below the diagnostic threshold at post-treatment on the CGI-TTM, and slightly more than 60% remained so at one-year follow-up.
• Overall, patients improved equally well during treatment and it was not until the one-year follow-up that differences in treatment response came to the fore.

Abstract
Background
This study sought to investigate the effectiveness of group treatment for trichotillomania (TTM) in ordinary clinical settings. Treatment consisted of a combination of habit reversal training (HRT) and acceptance and commitment treatment (ACT). Both short- and long-term effects were explored, as well as individual change trajectories.

Methods
The sample consist of fifty-three patients with TTM. Treatment outcomes were evaluated at post-treatment and at one-year follow-up using self-report questionnaires (Massachusetts General Hospital Hair Pulling Scale, MGH-HS), structured clinical interviews (National Institute of Mental Health Trichotillomania Severity Scale, NIMH-TSS), and the Clinical Global Impression scale for TTM (CGI-TTM).

Results
Analyses by mixed models for repeated measurements yielded a statistically significant effect of time (p<.001) for all outcome measures. There were large effect sizes, ranging from 1.76 to 2.33 at post-treatment and from 1.03 to 1.43 at one-year follow-up. Nearly 90% of patients scored below the diagnostic threshold at post-treatment on the CGI-TTM, and slightly more than 60% remained so at one-year follow-up. There were large and statistically significant differences in the change profiles across the patients from post-treatment to one year follow-up.

Conclusions
ACT-enhanced behavior therapy in a group format seems efficient for reducing symptoms of trichotillomania.

Keywords
Trichotillomania, Hair pulling disorder, Habit reversal training, ACT, Group therapy

1. Introduction

Trichotillomania (TTM) is characterized by repetitive pulling or removing of one’s own hair, typically from the scalp, eyelashes and eyebrows, resulting in hair loss. In the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), TTM was reclassified from an impulse control disorder to obsessive and compulsive related disorders (American Psychiatric Association, 2013). TTM affects between .6% and 3.5% of the general population (Christenson et al., 1991, Duke et al., 2009), has an age of onset ranging from 13 to 16 years, and follows a chronic course associated with considerable psychosocial impairment (Flessner et al., 2009, Stemberger et al., 2000). Some TTM patients spend several hours a day engaged in hair pulling (Stanley, Swann, Bowers, Davis, & Taylor, 1992). Awareness of the disorder,
even among health professionals, is low, and access to evidence-based treatment is limited (Marcks et al., 2006, Woods et al., 2006).

**Behavioral therapy (BT)** is the intervention with strongest evidence of efficacy for adults and children suffering from TTM (Bloch et al., 2007, Franklin et al., 2011, McGuire et al., 2014, Snorrason et al., 2015). Most studies examining BT have investigated habit reversal training (HRT), an approach in which patients are taught to be aware of their hair pulling and to perform behaviors incompatible with pulling (e.g., clenching their fists, folding their arms or hands) when pulling, pulling-related behaviors or pulling urges occur (Morris, Zickgraf, Dingfelder, & Franklin, 2013). Overall BT has been found to be effective for TTM with *effect sizes* (ES) ranging from .66 to 2.28 in several controlled trials (Azrin et al., 1980, Keijsers et al., 2006, Ninan et al., 2000, van Minnen et al., 2003). However, few patients achieve full abstinence from hair pulling, and relapse is a significant problem (Keijsers et al., 2006, Lerner et al., 1998). The study with the longest available follow-up data suggests 70% of the responders relapse within two years after HRT (Keijsers et al. (2006).

In an effort to improve both the short- and long-term impact of BT, several researchers have added treatment elements that focus not only on overt behaviors but also on associated thoughts and feelings. Enhancements to BT have included cognitive restructuring (Diefenbach, Tolin, Hannan, Maltby, & Crocetto, 2006), dialectical behavioral therapy (DBT- Keuthen et al., 2011; Keuthen et al., 2012), and acceptance and commitment therapy (ACT- Woods et al., 2006Twohig & Woods, 2004). A recent *meta-analysis* by McGuire et al. (2014) found greater effect sizes from both DBT- and ACT-enhanced behavioral treatment protocols compared to standard BT. Maintenance of treatment effects was also better with ACT-or DBT-enhanced BT approaches at 3 and 6 months post-treatment. However, there is a lack of follow-up data on enhanced BT beyond 6 months (Woods et al., 2006, Keuthen et al., 2011, Keuthen et al., 2012).

The efficacy of a protocol combining HRT with ACT has been demonstrated by Woods and colleagues in one RCT showing that ACT-enhanced HRT was superior to a waitlist control (Woods, Wetterneck, & Flessner, 2006). Symptom improvements were maintained at 3-month follow-up. The acceptance and commitment therapy (ACT) model postulates that hair pulling is maintained, in part, because patients exhibit heightened degrees of “experiential avoidance”, or an unwillingness to experience certain aversive thoughts, feelings or urges (Hayes & Wilson, 1994; Hayes, Wilson, Gifford, Follette, & Strosahl, 1996). ACT procedures focus on teaching the client skills to step out of the struggle with these internal experiences. Through use of metaphors and experiential exercises, patients learn to view urges, thoughts or feelings as stimuli to be observed rather than acted on. Patients learn that attempts to control these experiences have a long term cost, but by taking an accepting stance toward aversive thoughts, feelings and urges, the patients learn to instead engage in values-driven activities and behaviors. Ultimately, aversive experiences, when present, become less problematic.

Beyond the symptoms of pulling, embarrassment, shame and fear of social rejection are common in persons with TTM (G.J. Diefenbach et al., 2005, Stemberger et al., 2000). For this reason, group treatment may provide an advantage over individual treatment. Groups may relieve stigma and social alienation by providing support and a means of identifying with affected peers. According to a recent conceptual review of shame in obsessive-compulsive and related disorders (Weingarden & Renshaw,
third wave therapies such as ACT and DBT were recommended to reduce shame and enhance treatment effectiveness. The same authors also recommended therapy in a group setting in order to gain relief from shame and increase treatment effects in obsessive-compulsive disorders and related disorders (e.g., TTM).

Despite promising evidence in studies of group BT (Mouton and Stanley, 1996, Diefenbach et al., 2006, Toledo et al., 2015), there is preliminary evidence that group BT in the long-term is not superior to an active control (Diefenbach et al., 2006). In the one study, to examine the long-term impact of group BT, Diefenbach et al. (2006) compared BT to supportive therapy, provided over 8 weeks. Although group behavior therapy (ES=1.39) was superior to supportive group treatment (ES=.45) at post-treatment, treatment gains for the majority of participants were lost after six months. At post-treatment, 66.7% of the patients treated with BT were much or very much improved on the clinical assessments provided by the CGI, which was reduced to 33.3% after 6 months.

Most of the aforementioned treatment studies found larger standard deviations of outcome measures at follow-up as compared with standard deviations at baseline, indicating that treatment response differs from individual to individual. Testing these differences statistically is important from a clinical point of view as this constitutes the prerequisite of the search for predictors of treatment outcome and be helpful in capturing patients who may not respond well to treatment.

The majority of previous individual and group treatments for TTM have been conducted at highly specialized university centers with treatments provided by highly skilled expert providers. Thus, concerns have been raised regarding the ability to generalize from university clinics to ordinary clinical settings. This study of ACT-enhanced BT in a group setting for TTM is the first of its kind investigating this specific treatment modality. As far as we know this is also the first study to provide follow up data beyond 6 months for any program of enhanced BT or group BT for TTM. The primary aim was to investigate the effectiveness of ACT-enhanced group BT using a group-format in ordinary specialist services. In accordance with other studies of enhanced BT, we expected that there would be significant improvements in TTM symptoms from baseline to one-year follow-up. A secondary aim was to examine clinical change trajectories at an individual level. We anticipated that there would be significant differences in response rates across patients from pre- to post-treatment and from post-treatment to one-year follow-up.

2. Method

2.1. Participants

Fifty-three patients (4 men, 49 women) from three different treatment sites in Norway were included in the study. Participants ranged in age from 15 to 65. Six participants were under the age of 18. The three treatment sites recruited varying numbers of participants; an outpatient clinic focused on anxiety disorders in Oslo (n=35); a general community-based outpatient clinic in Kristiansand (n=6); and an OCD-specific treatment unit in Trondheim (n=12). Between 2013 and 2014, subjects were recruited via referrals from mental health outpatient clinics, student health care centers, letters to general practitioners and self-referrals from newspaper and web-based advertisements. Inclusion criteria
included a diagnosis of TTM obtained via a structured diagnostic interview for TTM (Rothbaum & Ninan, 1994) and a score ≥4 on the Clinical Global Impression Scale for TTM (G. Diefenbach et al., 2005, Guy, 1976), indicating at least moderate symptom severity. Exclusion criteria included the presence of schizophrenia spectrum disorders; substance addiction; severe antisocial, schizotypal, or paranoid personality disorder (PD); severe eating disorder; and ongoing suicidal ideation. Patients with bipolar disorder and ADHD were included if their symptoms were acceptably regulated. Characteristics of the participants who attended at least one session are presented in Table 1. Twenty patients reported using the following medications at intake: SSRIs (n=8), lamotrigine (n=2), methylphenidate (n=3), antipsychotic (n=1), mirtazapine (n=5), anxiolytic (n=1). Medication stability was not measured, but the patients were asked to maintain a steady dosage of medication throughout the active treatment phase of the study. Written informed consent was obtained from the entire sample. The project was approved by the Regional Ethics Committee for Medical Research with human subjects.

Table 1. Pre-treatment status on demographic and diagnostic variables for intention to treat (N=53).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>31.0 (11.7)</td>
</tr>
<tr>
<td>Onset age of TTM</td>
<td>12.9 (6.2)</td>
</tr>
<tr>
<td>Duration of TTM in years</td>
<td>17.4 (12.4)</td>
</tr>
<tr>
<td>% (n)</td>
<td></td>
</tr>
<tr>
<td>Female gender</td>
<td>92.5 (49)</td>
</tr>
<tr>
<td>Single</td>
<td>52.9 (28)</td>
</tr>
<tr>
<td>Married/cohabiting</td>
<td>47.2 (25)</td>
</tr>
<tr>
<td>Employed</td>
<td>58.4 (31)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>15.1 (8)</td>
</tr>
<tr>
<td>Student</td>
<td>26.4 (14)</td>
</tr>
<tr>
<td>Using medication</td>
<td>28.3 (15)</td>
</tr>
<tr>
<td>Any depressive disorder</td>
<td>66.7 (34)</td>
</tr>
<tr>
<td>Any anxiety disorder</td>
<td>80.4 (41)</td>
</tr>
<tr>
<td>Eating disorder</td>
<td>18 (9)</td>
</tr>
<tr>
<td>Alcohol/abuse</td>
<td>5.8 (3)</td>
</tr>
<tr>
<td>Other diagnosis</td>
<td>17.6 (9)</td>
</tr>
<tr>
<td>No other Axis I disorder</td>
<td>11.4 (6)</td>
</tr>
</tbody>
</table>

2.2. Procedure

Potential participants were screened for eligibility during a short telephone interview. A total of 71 participants passed the phone screen and were admitted for face-to-face clinical assessment. During the first face-to-face visit, participants were administered the Anxiety Disorder Interview Schedule (ADIS-IV) (Brown, Di Nardo, & Barlow, 1994), Structured Clinical Interview for DSM-IV Personality Disorders (SCID-II; First, Gibbon, Spitzer, & Benjamin, 1997); and the TTM DSM-IV diagnostic interview (Rothbaum & Ninan, 1994). The National Institute of Mental Health Trichotillomania Severity Scale (NIMH-TSS) was administered at baseline, at post-treatment, and at one-year follow-up. All interviews
and assessments were conducted by independent evaluators (BH or EMM). Clinical Global Impression for TTM (CGI-TTM) was also rated on these three points in time. Of the 71 patients admitted to the structured diagnostic assessment, 18 participants were excluded for the following reasons: did not attend the assessment (n=6), dropped out during the assessment (n=3), did not meet the inclusion criteria (n=6), offered individual therapy because of schedule problems (n=1) and treatment was postponed due to practical issues (n=2).

A total of 53 participants began treatment. Four patients dropped out during treatment. Drop-outs occurred for the following reasons: social phobia/fear for group sessions (n=1); physical illness (n=1); serious family-related distress (n=1) and work (n=1). A total of 49 patients completed treatment and were included in the treatment completer analyses. A patient had to attend at least six sessions to be defined as a completer. The completers attended on average 86% of the sessions (range 71–100%). There were no significant differences in session attendance across the three sites. The number of the patients who attended booster sessions were as follows: after 2 months (n=40), after 4 months (n=35) and after 6 months (n=29). Participants were also given a package containing self-report questionnaires (described below) at baseline, at the end of the treatment and at one-year follow-up.

2.3. Measures

2.3.1. Massachusetts General Hospital Hairpulling Scale (MGH-HS)

The MGH-HS (Keuthen et al., 1995) is a seven-item self-report measure, which assesses hair pulling severity. Each response is scored on a 5-point Likert-type scale (0=low severity; 4=high severity) with a total score ranging from 0 to 28. On the MGH-HS, higher numbers represent greater severity. The MGH-HS has generally demonstrated acceptable test-retest reliability, adequate sensitivity to change when correlated with other measures of TTM and good convergent and discriminant validity with respect to other measures of psychological functioning (Keuthen et al., 1995; O’Sullivan et al., 1995). Cronbach’s alpha in our dataset for the MGH-HS at the three measurement occasions was .88, .88, and .94, indicating good internal consistency.

2.4. National Institute of Mental Health Trichotillomania Severity Scale (NIMH-TSS)

The NIMH-TSS (Swedo et al., 1989) is a clinician completed rating scale measuring overall impairment produced by the time spent pulling or concealing damage, ability to control pulling, severity of alopecia, interference, and incapacitation caused by the pulling. NIMH-TSS total scores range from 0 to 25. Studies have shown mixed psychometric properties for NIMH-TSS (G. Diefenbach et al., 2005, Stanley et al., 1999). Cronbach’s alpha in our dataset sample for NIMH-TSS at the three measurement occasions was low at baseline (.52) but high at post-treatment (.81) and at one-year follow-up (.86). Inter-item correlation (r=.016) was lowest for items 1 and 5 (“Time spend on hair pulling last week” and “How much are you bothered by this compulsion/habit?”).
2.5. Clinical Global Impression Scale for TTM (CGI-TTM)

The CGI modified for TTM (Guy, 1976, Diefenbach et al., 2005) requires that clinicians score the severity of TTM symptoms on the basis of a global impression of symptom severity ranging from 0 to 7: (0= not assessed; 1= normal, not a disorder; 2= sub-clinical disorder; 3= mild disorder; 4= moderate disorder; 5= moderate /severe disorder; 6= Severe disorder; 7= very severe disorder.

2.6. Treatments

Treatment was conducted in groups of 3–7 participants with 2 therapists facilitating each group. Some groups had fewer than 6–7 patients in order to prevent too long of a wait. Ten groups were conducted in the study period. Each group met weekly for 3 h per session over 10 weeks. The treatment protocol followed the ACT-enhanced BT manual (Woods & Twohig, 2008). The Norwegian translation of the manual was published in April 2013. Similar to the English version, the Norwegian translation was organized for an individual format. In the current study, this individual format was wholly adopted as a group-based treatment with one modification, which included one additional meeting for family members between the third and fourth sessions. It is important to note that this modification only occurred at the Oslo treatment site (n=35), as a site-specific enhancement.

ACT-enhanced BT is a combination of traditional BT procedures as well as acceptance and commitment therapy (ACT). BT contains two components: habit reversal training (HRT, i.e., awareness training and competing response training) and stimulus control (SC). The purpose of awareness training is to help the client recognize and react to pulling. Awareness training involves describing the pulling as well as describing the sensations and situations that trigger pulling. Competing response training is the most central component in HRT. A competing response is a behavior that is physically incompatible with pulling and is something that the client can do easily in almost any situation (e.g., folding the arms together, folding the fingers around the thumbs, putting one’s hands in pockets). Stimulus control involves reducing or eliminating concrete triggers of hair pulling in the environment (e.g., removing a chair with armrests, choosing a different place to sit when watching TV, dampening the light in the bathroom or wearing band aids on the finger tips involved in hair pulling to diminish sensory stimuli) (Woods & Twohig, 2008).

ACT components are designed to establish private events as stimuli that do not need to be acted upon. Instead, the client is encouraged to behave in ways consistent with their self-identified values (e.g., being a good mother/father) and learn skills to step out of or accept the urge to pull, rather than fighting against the urge. The therapist helps the client to identify what is important to him/her and how fighting against the pulling urges have held the client back from engaging in activities that are driven by his/her values (as opposed to engaging in activities that are driven by urges). A brief description of the intervention is provided below:

Session 1: An overview of treatment and TTM-related psycho-education was provided.
Session 2: Implementation of HRT and SC procedures.
Sessions 3–8: Reviews of HRT and SC procedures as well as an initiation and integration of ACT-based components.
Session 9–10: Reviews of previous material and implementation of relapse prevention techniques.

An important component in this treatment manual was the systematic use of a workbook, which the participants received at the beginning of the treatment. The workbook contained educational materials, forms for self-completed behavioral analyses, and forms for recording behavioral exercises. At the start of every session, clients were asked to submit weekly assessments of their hair-pulling activity. To reinforce the clients' progress, data from these measures were plotted on a graph and reviewed weekly with the client in the presence of other group members. The participants were also invited for three group-based booster sessions after 2, 4 and 6 months. The content in the booster sessions included reviews of HRT and ACT procedures, along with progress updates for each participant.

2.7. Therapists

Therapy was provided by 8 different therapists, including 6 psychologists, 1 psychiatric nurse and 1 psychiatrist. All therapists attended a two-day pre-study training seminar with the author of the manual, Douglas W. Woods. All, but one therapist had limited clinical experience treating TTM prior to the study. During the study the therapists attended several supervision meetings. Most assessment sessions were videotaped.

Adherence and integrity checklists were developed by three of the authors of the study (DWW, BH and PAV) and closely followed the manual. Three of the 10 sessions from each group were rated to cover both major active treatment components (BT & ACT). The videotapes from one group held in Trondheim and two groups held in Oslo were viewed by two pairs of master degree psychology students who rated sessions as a pair to achieve consensus ratings. No formal inter-rater reliability checks were obtained, but when consensus could not be achieved, the students met with an experienced clinical psychologist (PAV) to reach a consensus rating. Adherence scales involved a checklist of topics that should have been covered in each session (0= not covered, 1= covered). The average adherence to the manual from the groups ranged from 82% to 92% and was considered satisfactory. An overall rating of session quality (integrity) was evaluated on a four-point scale (1= poor, 2= moderate, 3= good, and 4= very good). These ratings were subjective and meant to assess the perceived quality of the BT and ACT interventions used, such as the clarity of examples and perceived quality of the discussion obtained with the patients. The average quality of the three sessions was 3.0 in Trondheim and 2.7 for both groups in Oslo. These ratings were considered satisfactory.

2.8. Statistical analyses

Mixed models procedures in SPSS were used to analyze the longitudinal data, with MGH-HS, NIMH-TSS, and CGI-TTM as outcome (dependent) variables. Log likelihood (LLH) was used to evaluate model fit. For the linear model, the best fit was for the model with random intercept, random slope, and
autoregressive covariance matrix for residuals (Fitzmauritz, Laird, & Ware, 2004). This was true for all outcome variables. Inspection of scatter plots indicated that a linear spline model might fit the data better than a simple linear model. Indeed, fit improved considerably by including a linear spline, with the knot at post-treatment. Among all models being tested, the best fit was for the spline model with random intercepts and random slopes for both slopes (from pre-treatment to post-treatment and from post-treatment to one-year follow-up). This model was used in testing individual change trajectories. Identification of outliers/unusual change profiles was accomplished by a Shapiro-Wilk test and Kolmogorov-Smirnov test for normality.

Differences in change rates between subgroups (treatment drop-outs versus treatment completers; patients with missing MGH-HS versus patients without missing MGH-HS; patients from Oslo versus patients from Kristiansand/Trondheim) were tested by including two-way interactions in the mixed models (time by group).

To facilitate comparisons with previous studies, Cohen's d for one-sample repeated measures was used to compute effect sizes (Cohen, 1988). The pooled standard deviation was used as the denominator in this study.

2.9. Missing data

At post-treatment, NIMH-TSS and MGH-HS data were available for 48 and 47 patients, respectively (91% and 89% response rate). At the one-year follow-up, NIMH-TSS and MGH-HS data were available for 51 and 44 patients, respectively (96.2% and 83% response rate). Data were imputed by predicted scores at one-year follow-up, obtained by linear regression with MGH-HS as the predictor variable and NIMH-TSS as the dependent variable, or vice versa. Data at post-treatment were not imputed. For the two patients who had neither NIMH-TSS nor MGH-HS at follow-up, pre-treatment data were carried forward.

2.10. Reliability analyses

Fourteen NIMH-TSS interviews conducted by the last author (BH) were also rated by the third author (EMM). Inter-rater reliability (correlation coefficient) was .871 for CGI-TTM (p<.001) and .906 for NIMH-TSS (p<.001), which indicates good inter-rater reliability. Closer inspection revealed that most disagreement occurred on the sixth item of the NIMH-TSS ("How much does hair pulling interfere with your daily life?"). Indeed, deleting this item resulted in a correlation coefficient of .985. Moreover, IRR was lower at baseline than at one-year follow-up (r=.411 versus r=.969, respectively). This discrepancy can be explained by the achievement of better inter-rater agreement with respect to item 6 during the course of the project (we agreed upon rating functional impairment only). It should be mentioned that the first rater, of whom the results were used in this study, had somewhat lower means of NIMH-TSS than the second rater (15.3 versus 16.7, respectively, p=.025).

The correlation between MGH-HS and NIMH-TSS was lower at pre-treatment (r=.473, p=.000) than at post-treatment (r=.613, p=.000) or one-year follow-up (r=.853, p=.000). Inspection of the scatter plots revealed that at pre-treatment, NIMH-TSS scores were relatively high compared with MGH-HS scores.
This effect was not present at post-treatment or at the one-year follow-up. This may, in part, be explained by the fact that at baseline, the MGH-HS was filled in between the first and second assessment session, whereas the NIMH-TSS was assessed at the first session, taking into account TTM symptoms the week before. For many patients, this was the first time in their life that they had talked about TTM with a professional, which may have given them temporary relief of symptoms, and thus, lower MGH-HS scores. Another likely explanation is related to the very effect of treatment. In BT, patients are not only being taught to be more aware of their hair pulling behavior but also of behavior that is associated with hair pulling, such as social avoidance behavior and checking behavior (e.g., checking whether bald spots might be visible for others). Thus, more awareness of hair pulling related behavior may have resulted in higher MHG-HS scores.

3. Results

3.1. Intention-to-treat-analyses

Table 2 displays the means, standard deviations and effect sizes for intention-to-treat patients for the three assessment points for MGH-HS, NIMH-TSS and CGI-TTM. There were large effect sizes for the outcome measures, ranging from 1.76 to 2.33 at post-treatment and from 1.03 to 1.43 at the one-year follow-up (Cohen, 1988). Mixed models analyses revealed that there were significant changes for all outcome variables from baseline to the one-year follow-up. In more detail, MGH-HS scores were reduced by 6.0 points from baseline to the one-year follow-up (t=- 5.4, p=.000); NIMH-TSS scores were reduced by 7.9 points (t=- 8.6, p=.000); and CGI-TTM scores by 1.7 points (t=- 8.6, p=.000) (Fig. 1).

Table 2. Number of patients (N), mean (M) and standard deviations (SD), and effect sizes (ES) for MGH-HS, NIMH-TSS and CGI-TTM at pre-treatment, post-treatment and one-year follow-up (intention to treat).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGH-HS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-treatment</td>
<td>52</td>
<td>17.86</td>
<td>5.02</td>
<td>1.76</td>
</tr>
<tr>
<td>Post-treatment</td>
<td>53</td>
<td>8.58</td>
<td>5.52</td>
<td>1.03</td>
</tr>
<tr>
<td>One year F.U.</td>
<td>52</td>
<td>11.7</td>
<td>6.83</td>
<td></td>
</tr>
<tr>
<td>NIMH-TSS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-treatment</td>
<td>53</td>
<td>16.46</td>
<td>4.15</td>
<td></td>
</tr>
<tr>
<td>Post-treatment</td>
<td>53</td>
<td>6.23</td>
<td>5.50</td>
<td>2.10</td>
</tr>
<tr>
<td>One year F.U</td>
<td>53</td>
<td>8.52</td>
<td>6.21</td>
<td>1.51</td>
</tr>
<tr>
<td>CGI-TTM</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-treatment</td>
<td>53</td>
<td>4.73</td>
<td>.73</td>
<td></td>
</tr>
<tr>
<td>Post-treatment</td>
<td>53</td>
<td>2.32</td>
<td>1.27</td>
<td>2.33</td>
</tr>
<tr>
<td>One year F.U</td>
<td>53</td>
<td>3.07</td>
<td>1.47</td>
<td>1.43</td>
</tr>
</tbody>
</table>

Note. MGH-HS=Massachusetts General Hospital Hair pulling Scale; NIMH-TSS= National Institute of Mental Health Trichotillomania Severity Scale; CGI-TTM= Clinical Global Impression Scale for TTM
One may argue that positive change, as estimated by the linear model, may be due to the strong post-treatment effect. Therefore, we conducted mixed models analyses in which the post-treatment data were deleted. These analyses still gave significant change parameters (p<.001) for all three outcome variables (decreases in symptoms).

3.2. Clinical improvement analyses

Clinical improvement was assessed by comparing CGI-TTM scores before and after treatment. At pre-treatment, all patients had a CGI-TTM score between 4 and 6. At post-treatment, 87.5% would not have met the study entry criteria (score below 4), though this was reduced to 62.8% at the one-year follow-up. At post-treatment 34% (n=18) had a cut-off score on the MGH-HS of 6 or below, which dropped to 22.6% (n=12) at the one-year follow up. MGH-HS scores of 6 or below are defined as a clinically significant response according to Diefenbach et al. (2006).

3.3. Individual change trajectories

To assess individual change trajectories, a linear spline model was taken into consideration first. A linear spline model implies that change rates are computed separately for the treatment phase (baseline to post-treatment) and the follow-up phase (post-treatment to one-year follow-up), within the same model. In our analyses, it appeared that there was a large decrease in symptoms during the ten-session treatment phase, which was significant at the p<.001 level for all variables (9.9 points on average for MGH-HS, t=−11.6; 11.1 points for NIMH-TSS, t=−15.1; and 2.6 points for CGI-TTM, t=−16.5). However, from post-treatment to one-year follow-up, TTM symptoms showed a slight, but significant,
increase (p<.001) for all outcome measures (3.9 point increase on average for MGH-HS, t=4.0; 3.2 points for NIMH-TSS, t=3.9; and .95 for CGI-TTM, t=5.3) (Fig. 2).

Fig. 2. Clinical course for NIMH-TSS; linear spline model. Note: The mean change rate for NIMH-TSS from baseline to post-treatment was 11.1 points improvement (SD=.74, p<.001). From post-treatment to one-year follow-up, there was a 3.2 points increase in symptoms (SD=.83).

The next step in the procedure is the analysis of random effects. A “full-blown” mixed model analysis implies that both fixed and random effects are included in the same model. A “random slope at the individual level” implies that for each patient, a separate change rate is estimated, i.e., an individual regression coefficient. Accordingly, each patient deviates from the “main slope” (the fixed effect) to a certain degree, and the variance of these deviation scores (the random effects) can be computed. In a spline model, these slope variances are computed separately for the two phases of this study, i.e., the baseline to post-treatment phase, and the post-treatment to one-year follow-up phase.

In our sample, the variances of random slopes were significant for the post-treatment phase (σ=13.8, p=.029 for MGH-HS; σ=18.1, p=.035 for NIMH-TSS; and σ=.97 p=.003 for CGI-TTM). However, for the treatment phase, the variances of random slopes were not significant for MGH-HS and NIMH-TS (σ=5.2, p=.263 for MGH-HS; σ=6.2, p=.104 for NIMH-TSS) and were substantially smaller for CGI-TTM (σ=.58, p=.006). It should be noted that during treatment, although not significant, there was still some variance in the random slopes. These findings indicate that the patients had quite similar, but not identical change rates during treatment and very different change rates after treatment, i.e., some patients had additional improvement, whereas others relapsed.
3.4. Control analyses

Three out of four treatment drop-out patients participated in the one-year follow-up examination (only clinical interview, i.e., NIMH-TSS). Compared with treatment completers, drop-out patients had lower levels of TTM symptoms at baseline (significantly different intercept, \( p=.024 \)) but did not have significantly different change rates (parallel change, \( p=.832 \)). Thus, the overall results were probably not influenced by having included drop-out patients in the intention-to-treat analyses.

Another potential bias concerns the imputation of missing MGH-HS data at the one-year follow-up. However, mixed models analyses in which the eight patients with missing MGH-HS were excluded gave highly similar or identical results. The only difference was a slightly better change rate for the MGH-HS (3.3 instead of 3.0 per time unit). For the linear spline model, change rates were identical. The validity of these findings was confirmed by way of mixed models analyses in which patients with missing MGH-HS were recoded as a dummy variable and included as a two-way interaction in the model with NIMH-TSS as the dependent variable. There was no significant effect for this interaction and very small effect sizes: \( F=.079, p=.779 \) for the linear model and \( F=.82 \ p=.368 \) for the linear spline model (follow-up phase). In addition, at the one-year follow-up there were no differences in mean NIMH-TSS between the two groups (independent sample \( t \)-test, \( t=.87, p=.39 \)).

To investigate whether patients at the Oslo clinic had better outcomes than the other patients, treatment unit was recoded as a dummy variable (Oslo versus Trondheim/Kristiansand) and included in the analyses by way of a two-way interaction (treatment unit by time). This interaction was not significant for any outcome variable. This interaction was also not significant when Trondheim patients were compared with Oslo/Kristiansand patients. Table 3 displays pre-treatment scores for MGH-HS, NIMH-TSS and CGI-TTM at each treatment site.

Table 3. Number of patients (N) at each site, mean (M) and standard deviations (SD) for MGH-HS, NIMH-TSS and CGI-TTM at pre-treatment (intention to treat).

<table>
<thead>
<tr>
<th>Site 1 (n=35)</th>
<th>Site 2 (n=6)</th>
<th>Site 3 (n=12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGH-HS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-treatment</td>
<td>18.0* (.76)</td>
<td>13.8 (3.09)</td>
</tr>
<tr>
<td>NIMH-TSS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-treatment</td>
<td>15.50 (.66)</td>
<td>16.08 (1.81)</td>
</tr>
<tr>
<td>CGI-TTM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-treatment</td>
<td>4.64 (.13)</td>
<td>4.83 (.30)</td>
</tr>
</tbody>
</table>

Note: Site 1= Oslo, Site 2= Kristiansand, Site 3= Trondheim, MGH-HS=Massachusetts General Hospital Hair pulling Scale; NIMH-TSS= National Institute of Mental Health Trichotillomania Severity Scale; CGI-TTM= Clinical Global Impression Scale for TTM, * = n=34 due to missing.
3.5. Residual analyses/outliers

Shapiro-Wilk tests for normality were significant for all models, except for the NIMH-TSS spline model. Kolmogorov-Smirnov tests were not significant. Inspection of outlier plots revealed that one patient had a deviant change profile for the NIMH-TSS, and four patients had deviant change profiles for the MGH-HS. Mixed models analyses of the MGH-HPS data without these patients gave almost identical change rate estimates and normalized Q-Q plots. However, from pre-treatment to post-treatment, the slope variance became larger (from 5.2 to 10.2) and reached statistical significance (p=.040). From post-treatment to one-year follow-up, slope variance decreased from 18.1 to 13.8 (p=.065).

4. Discussion

The main purpose of the current study was to test the effectiveness of group ACT-enhanced BT for Norwegian TTM patients. The results showed significant decreases in MGH-HS, NIMH-TSS and CGI-TTM scores from pre-treatment to the one-year follow-up. The drop-out rate was low (7.5%), and the average attendance for each session was high (86%), indicating that the treatment was highly acceptable. Although the results revealed a slight, but significant, worsening in symptoms from post-treatment to the one-year follow-up, the effect sizes of the MGH-HS, NIMH-TSS and CGI-TTM were still high (i.e., 1.01, 1.51 and 1.4) at the one-year follow-up.

Contrary to expectations, individual response rates did not differ significantly during the treatment for two of the three outcome variables (i.e. NIMH-TSS and MGH-HS). However, from post-treatment to the one-year follow-up, there were large and significant differences in the change trajectories across patients, indicating that some patients had additional improvement, whereas others relapsed. Thus, response during treatment might not be a good indicator of treatment response in the long run.

These results are consistent with prior meta-analyses (Bloch et al., 2007, McGuire et al., 2014, Snorrason et al., 2015, Slikboer et al., 2015) showing that a range of different forms of behavior treatments are efficacious for TTM. However, compared with other group studies, our results indicate that ACT-enhanced BT might be superior to BT or CBT, as our one-year follow-up effect size (ES=1.76) on the MGH-HS was larger than the pre-post effect size (ES=1.02) reported by Diefenbach et al. (2006) at 6-month follow-up for the BT condition. Moreover, our post-treatment effect size was somewhat larger than the 1.35 effect size of the CBT condition in the study of Toledo et al. (2015). It should be noted that Diefenbach and colleagues reported an effect size of 1.39, but when using pooled standard deviations, as we did in our study, we computed an effect size of 1.02 for their data.

The maintenance of improvement at the one-year follow-up found in the present study also indicates more desirable results than individual CBT/BT (Keijzers et al., 2006, Lerner et al., 1998). On the other hand, Keuthen et al. (2012) found effect sizes that exceeded our effect sizes in a individual DBT-enhanced BT study. Moreover, as opposed to our results, improvements were largely maintained throughout the 6-month follow-up. There are several possible explanations for these differences. Individual therapy may be an advantage over group treatment. Furthermore, as noted by Keuthen et al. (2012), DBT is less conceptually abstract than ACT and offers more specific skill instruction for
specific targets. DBT might also be easier to implement and combine with BT since it is not necessary for experienced cognitive-behavioral therapists to learn new theoretical backgrounds. On the other side, differences between our study and Keuthen’s study could also be explained by the longer follow-up period used in our study, increasing the probability of relapse. Sample bias may be another explanation, as the study of Keuthen (2012) had more strict exclusion rules, excluding patients with ADHD, bipolar disorder type I, and suicidality, making factors that may increase the likelihood of relapse lower in the Keuthen et al. study.

Although the results are promising, the open trial design does not allow us to conclude whether the specific techniques (e.g., ACT, HRT) or more generic components (e.g., social support, alliance, group setting) were responsible for the symptom improvements. Interestingly, there is some evidence that the group setting may be important for the treatment outcome. In the study by Diefenbach et al. (2006), both group behavior therapy and supportive group therapy significantly reduced TTM symptoms.

Our clinical impression is that social support from peers is particularly useful in managing underlying negative emotions and beliefs in TTM. Notably, treatment in groups seems to be particularly helpful in decreasing a sense of shame related to TTM. Reducing shame and providing a sense of normalization were mentioned by participants as two of the most important aspects of therapy.

4.1. Limitations

One central limitation of the present study was the open trial design with the lack of a control condition. Nevertheless, it is unlikely that our results are due to non-specific factors and participant maturation given the chronicity of symptoms in our sample (duration of TTM is approximately 17 years) and given that our sample out-performed the supportive therapy benchmark of Diefenbach et al. (2006). Another limitation is the low correlation between the MGH-HS and NIMH-TSS at pre-treatment. As explained in the method session, we believe that the clinician ratings (NIMH-TSS) at baseline provide a more valid picture of patients’ symptom level prior to treatment than the MGH-HS and thus, the effect sizes of the MGH-HS might be too small. Poor internal consistency of the NIMH-TSS at baseline is also a limitation (alpha = .52). Though it is not ruled out that this might be due to low inter-rater reliability (IRR) at baseline, there are other potential explanations. It appeared that the poor internal consistency was mostly due to low inter-item correlation between two items, i.e., “Time spent on hair pulling last week” and “How much are you bothered by this compulsion/habit”. In the clinical interview, automatic pulling was explicitly taken into account and rated in this study. Before treatment, patients may spend much more time at automatic pulling than on focused pulling, to such an extent that there is no association between time spent on pulling and the degree patients are bothered by this habit. Thus, the low internal consistency of NIMH-TSS before treatment may also be explained by a lack of association between automatic hair pulling and the distress caused by this habit. Future research should focus on validating the NIMH-TSS with respect to the two types of pulling styles, i.e., automatic versus focused hair-pulling.

A note of caution is warranted concerning the effect sizes of our study. Pre-treatment standard deviations were quite small, especially for the NIMH-TSS (.73), which may have resulted in effect size
estimates that were artificially high. The low SD is partly due to selection bias, as only patients who at least were moderately ill (CGI-TTM score of 4) were included. However, even if the effect sizes were halved in the present study, effect sizes are still large and comparable with most published RCTs on both individual-and group treatments for TTM.

Furthermore, the treatment integrity was not ideal, but it was in our view acceptable considering that this was a naturalistic program evaluation. Only slightly above average adherence to the specific factors of the manual could actually be a strength, showing that the manual is generalizable to non-research settings. However, the significant reductions to TTM post-treatment suggest that perhaps other non-specific treatment factors, not measured by TTM ratings, compensate for mediocre integrity ratings.

Strengths of the current study include low drop-out rates, relatively large sample size, high rates of adherence, limited missing data, the use of both self-report and clinician-rated measures, assessments by independent raters, application of inter-rater reliability tests on the NIMH-TSS and CGI-TTM, and the use of formal fidelity analyses of the therapy sessions.

4.2. Implications for implementation

The majority of previous TTM studies have been conducted at specialized university clinics with treatment provided by highly skilled therapists. Critics have argued that controlled trials conducted in an academic setting bear little resemblance to community-based practice (Barlow et al., 1999, Shafran et al., 2009). According to Barlow et al. (1999), obstacles to dissemination may be related to a lack of effectiveness research exploring the value of CBT in real life contexts. The most important clinical implications for the present study may be that ACT-enhanced group BT for TTM appears to work well in non-university clinics, conducted by therapists with limited prior BT or ACT training. As outlined in the introduction, there is a striking gap between the demonstrated efficacy of CBT for TTM and other disorders and how poorly the treatment is disseminated worldwide. The current findings may encourage other therapists to implement group treatment in their units. Because group CBT is less costly and time-consuming, it may be an appropriate alternative to individual treatment. The main objection to delivering group therapy for TTM patients has been the difficulty of organizing groups because of low prevalence of TTM in clinical populations. Another objection has been that feelings of shame and stigma could be an important reason for avoiding groups. A study by Glazier, Wetterneck, Singh, and Williams (2015) found that across obsessive-compulsive and related disorders, the most frequently endorsed barrier to treatment was being “ashamed of my problems.” Overall it was not difficult to recruit enough patients in the current study. However, at the smallest site in the current study, it was somewhat problematic to assemble large enough groups within a reasonable time, which resulted in two groups with three patients.

In conclusion, this study indicates that ACT-enhanced group behavioral therapy leads to large improvements in TTM symptoms during treatment. Considering the high risk of relapse in TTM, the effect sizes were still large at the one-year follow-up. The individual patient trajectories indicate that response during the treatment has somewhat limited value in predicting long term outcomes. Future research should focus on other potential predictors of treatment outcome.
Acknowledgements

The authors would like to thank Gro Merete Grønvold Eilertsen, Anne Ruud, Torunn Grøtte and Lars Morså for serving as therapists in the project.

References

1 American Psychiatric Association (DSM-5)
29 S.G. Mouton, M.A. Stanley Habit reversal training for trichotillomania; A group approach Cognitive Behavioral Practice, 3 (1996), pp. 159-182


43 H. Weingarden, Renshaw Shame in the obsessive compulsive related disorders: A conceptual review *Journal of Affective Disorders*, 171 (2015), pp. 75-84
