Journal of Obsessive-Compulsive and Related Disorders 3 (2014) 345–353

Contents lists available at ScienceDirect

Journal of Obsessive-Compulsive and Related Disorders

journal homepage: www.elsevier.com/locate/jocrd

Predictors of relapse following treatment of trichotillomania

Martha J. Falkenstein *, Kate Rogers, Elizabeth J. Malloy, David A.F. Haaga

American University, United States

ARTICLE INFO

Article history:
Received 19 December 2013
Received in revised form
28 July 2014
Accepted 8 September 2014
Available online 2 October 2014

Keywords:
Trichotillomania
Hair-pulling disorder
Obsessive compulsive related disorder
Relapse
Maintenance

ABSTRACT

Objective: This study sought to identify predictors of relapse in a behavior therapy trial for trichotillomania (TTM), or hair-pulling disorder. Relapse is common after treatment for TTM, and only a few studies have examined what might predict relapse.
Method: Data was examined from a TTM treatment study with a stepped-care approach (Step 1. Web-based self-help; Step 2. Individual behavior therapy) ($N=60$). Implications of significant predictive relations were illustrated by constructing Probability of Treatment Benefit (PTB) charts (Lindhiem, Kolko, & Cheng, 2012), which quantify the probability of maintaining gains according to predictors of maintenance.
Results: Abstinence at the conclusion of treatment and lower TTM severity during initial response significantly predicted maintenance. Abstinence periods prior to treatment, residual urges after achieving abstinence, pre-treatment TTM severity, intrinsic motivation, and treatment compliance did not predict maintenance.
Conclusions: Post-treatment abstinence and lower TTM severity during initial response predicted maintenance. Replications of this research are needed to determine the usefulness of these possible predictors in identifying relapse-prone patients, with the aim of improving clinical decision-making and developing strategies to help these patients better maintain gains. This is the first TTM study to use PTB charts, which can help clarify the meaning of prognostic analyses.

1. Introduction

1.1. Background and significance

Trichotillomania (TTM), also called hair-pulling disorder, is characterized by compulsive hair pulling that results in hair loss, and is currently classified as an obsessive–compulsive related disorder. TTM is estimated to affect 1–2% of adolescents and young adults (American Psychiatric Association, 2013). It is associated with significant distress and impaired psychosocial functioning in many domains (Diefenbach, Reitman, & Williamson, 2000; Woods et al., 2006a).

Habit reversal training (HRT), a type of behavior therapy, is the intervention for adults with TTM with the most empirical support, according to a meta-analysis by Bloch and colleagues (Bloch et al., 2007) that indicated its superiority over pharmacotherapy. HRT consists of four interventions: 1) self-monitoring, to keep records of hair pulling, 2) awareness training, aimed at increasing awareness of hair pulling behaviors and situations that elicit pulling, 3) stimulus control, which involves techniques to prevent or interfere with pulling, 4) competing response training, the use of behaviors that are physically incompatible with pulling.

Efficacy studies of behavioral treatments for adults with TTM over the past decade and a half have been in a variety of formats and have generally had good response rates. Post-treatment response rates in studies of individual HRT have ranged from 64% to 100%; however few patients have achieved total abstinence from pulling (e.g., Lerner, Franklin, Meadows, Hembree, & Foa, 1998; Ninan, Rothbaum, Marsteller, Knight, & Eccard, 2000; van Minnen, Hoogduin, Keijsers, Hellenbrand, & Hendriks, 2003). In studies of group HRT, response rates have been more variable, ranging from 17% to 80% (Diefenbach, Tolin, Hannan, Maltby, & Crocetto, 2006; Mouton & Stanley, 1996). Recently, treatment studies have included emotion-regulation skills training with traditional HRT. In a study of acceptance and commitment therapy (ACT)-enhanced HRT, 66% of participants were defined as responders at post-treatment (Woods, Wetterneck, & Flessner, 2006b), and 80% were considered responders at post-treatment in a study of dialectical behavior therapy (DBT)-enhanced HRT (Keuthen et al., 2010).

Although treatment response rates have been satisfactory, failure to maintain gains from HRT is common, with 50-67% of treatment responders relapsing by the longest follow-up time point. More research with larger samples is needed to be confident in a precise

* Corresponding author. Tel.: +1 781 608 0896.
E-mail address: mf0136a@american.edu (M.J. Falkenstein).

http://dx.doi.org/10.1016/j.jocrd.2014.09.001
2211-3649/

estimate of relapse rates, but findings to date reflect that relapse appears to be a significant problem (see Table 1 for data from trials of behavior therapy for TTM; the table includes all published studies that provided information on the number of treatment responders who maintained gains in follow-up assessments). In the study of DBT-enhanced HRT, the relapse rate, 0%, is much lower than in other studies (Keuthen et al., 2011). Some participants did decline, though, from "full" to "partial" response status between follow-ups. This study could have yielded a lower relapse rate for one or more of several reasons: the addition of DBT to the typical HRT protocol; the increased number of sessions and extended period of time over which they occurred (23 weeks vs. 6–9 weeks); a definition of sustained response that may not be as stringent as in other trials (see Table 1), or sampling error. Regardless of the explanation for the lower relapse rate, several participants failed to fully maintain their initial gains, and in other studies many relapsed completely. Few studies have investigated individual differences that might predict relapse, which is important in order to identify relapse-prone patients and to better serve their needs during treatment.

1.2. Potential predictors of relapse in TTM

To guide the selection of relapse predictors to examine in this study, we turned mainly to conceptual and empirical work on addictive behaviors. Relapse is notoriously common after treatment of addictions. Moreover, it has been suggested that for at least some hair pullers, TTM might resemble an addiction; shared features include "(1) repetitive or compulsive engagement in the behavior despite adverse consequences; (2) diminished control over the problematic behavior; (3) an appetitive urge or craving state prior to engagement in the problematic behavior; and (4) a hedonic quality during the performance of the problematic behavior" (Grant, Odlaug, & Potenza, 2007, p. 81).

More specifically, we tested five possible predictors of relapse, identified as important in conceptual or empirical work on addictions and potentially applicable to TTM. First, we tested whether achieving complete abstinence [as opposed to reduction in frequency or severity but not to zero] from hair pulling would predict maintenance. Reviewing the literature on substance use, Miller and Carroll (2006) concluded that "changing a well-established pattern of drug use usually begins by interrupting the pattern to produce an initial period of abstinence" (p. 307). At least two studies have already reported data relevant to this hypothesis in TTM treatment. Lerner et al. (1998) observed that both of their study participants who had been abstinent at the completion of behavior therapy (BT) showed little symptom recurrence at long-term follow-up, whereas only two of the 11 non-abstaining participants were responders at follow-up. Keijsers et al. (2006) obtained similar results. Participants who had been abstinent from hair pulling at the end of treatment reported having fewer TTM symptoms at a two-year follow-up assessment than did those with only partial symptom reduction at post-treatment.

Second, we evaluated TTM symptom severity (separately at baseline and at the time of initial treatment response) as a predictor of maintenance. Studies of behavioral treatments for addiction show that "the effectiveness of these approaches tends to decrease with increasing severity of substance use and related problems" (Carroll & Rousnaville 2006, p. 236). Although the analysis did not clearly distinguish lack of initial response from relapse, it is suggestive that higher pre-treatment symptom severity predicted higher symptom severity at long-term follow-up ($M=3.75$ years post-treatment) in one TTM study (Lerner et al., 1998).

Third, history of relapses was tested as a predictor. Schachter (1982) hypothesized that having failed to maintain gains multiple times could actually be a favorable prognostic indicator for success on the index attempt. In particular, incremental learning may occur as the would-be behavior changer learns from each relapse what situations, moods, etc. are personally relevant triggers for relapse. However, it is also possible that the opposite direction of effect would be evident; that is, multiple previous relapses might serve as an indicator of someone who is highly vulnerable to relapse and likely to relapse again. In a recent smoking cessation study, participants with more prior relapses were indeed less likely to maintain abstinence after they initiated a quit attempt (Partos, Borland, Yong, Hyland, and Cummings, 2013). Thus, either direction of effect is plausible, and we are aware of no previous studies of this issue in TTM.

Fourth, we studied residual hair pulling urges as a predictor of relapse. In the framework described by Muraven and Baumeister (2000), residual urges reflect a vulnerability to relapse across a wide range of habitual behaviors. In the addictions, urge or craving has

Table 1
Long-term outcome in behavior therapy studies.

Study	Type of treatment	Definition of response to treatment (relapse defined as failure to maintain)	Responders at post-treatment	"Responders" who relapsed at follow-up
Mouton and Stanley (1996) ($n=5$)	6 weekly sessions of group HRT	Clinical significance on MTAI	80%	25% at 1 mo, 50% at 6-mos
Lerner et al. (1998) ($n=14$)	9 weekly sessions of individual HRT+cognitive therapy	$>$50% improvement on NIMH-TSS	86%	67% at M=3.75 yrs
Twohig and Woods (2004) ($n=6$)	7 weekly sessions of individual HRT+ACT	Descriptive comparisons of self-monitoring data, MGH-HPS, and photograph rating	67%	25% at 3 mos
Woods et al. (2006b) ($n=25$)	8 weekly and 2 biweekly sessions of individual HRT+ACT	MGH-HPS score of $\leq$12.74	66%	MGH-HS scores significantly higher at 3 mos than at post-treatment[a]
Diefenbach et al. (2006) ($n=12$)	8 weekly sessions of group HRT+cognitive therapy	MGH-HPS score of $\leq$6.0, following recovery of normal functioning guidelines (Jacobson & Truax, 1991)	17%	50% at both 3 and 6 mos
Keijsers et al. (2006) ($n=28$)	6 biweekly sessions of behavior therapy	$\geq$50% improvement on MGH-HPS	79%	63% at 2 yrs
Keuthen et al. (2011) ($n=10$)	11 weekly and 4 sessions over 3 months of individual HRT+DBT	Full responder: CGI$\leq$2 and $\geq$35% decrease in MGH-HPS scores	80% full responders	0% at both 3 and 6 mos
Keuthen et al. (2012) ($n=18$)	11 weekly and 4 booster sessions of individual HRT+DBT	Reference Keuthen et al. (2011)	61% full responders	MGH-HS scores significantly higher at both 3 and 6 mos than at post-treatment[a]

Note: MTAI=Minnesota Trichotillomania Assessment Inventory, NIMH-TSS=NIMH Trichotillomania Symptom Severity Scale, MGH-HPS=Massachusetts General Hospital Hairpulling Scale, CGI=Clinical Global Improvement Scale, ACT=Acceptance and commitment therapy, HRT=Habit reversal training, DBT=Dialectical behavior therapy.

[a] Maintenance information not provided on case-by-case basis.

been studied in relation to relapse often. For example, in a placebo-controlled clinical trial of nicotine gum, abstainers who reported stronger urges and cravings went on to relapse more quickly (Doherty, Kinnunen, Militello, & Garvey 1995). Some hair pullers likewise report strong urges or craving states prompting them to perform repetitive behaviors, resulting in a pleasurable or rewarding feeling during or after the problematic behavior (Brewer & Potenza, 2008). Others have urges to pull that are more like compulsions found in obsessive–compulsive disorder, typically aimed at reducing anxiety or other negative affective states (e.g., Diefenbach, Mouton-Odum, & Stanley, 2002). To date, it is unclear whether these urges, if still present after clinical improvement, predict relapse.

Fifth, we tested intrinsic motivation (measured pre-treatment as well as at the time of initial response) as a predictor of maintenance. According to Marlatt's Relapse Prevention model (e.g., Marlatt & Witkiewitz, 2005), the use of Motivational Interviewing methods to help substance users resolve ambivalence and see that continued use is likely to have consequences at odds with their own values and goals helps foster a sense of client autonomy and intrinsic motivation. Empirically, intrinsic motivation has correlated positively with maintenance of changes in health behaviors such as medication adherence and weight loss maintenance (Ryan & Deci, 2000), but to our knowledge no previous studies have examined this predictive relation in TTM treatment.

Finally, we tested treatment compliance as a predictor of maintenance, as it is believed that increased frequency of skills practice leads to more automatic usage in high-risk situations, which can result in better maintenance. This was demonstrated by Edelman and Chambless (1995) in a CBT trial for social phobia, in which increased homework adherence was associated with improved treatment outcome at 6-month follow-up.

The overarching objective of this study was to draw upon addiction theory and research in order to test potential predictors of the maintenance of response to behavior therapy for TTM. Predictors were tested in the context of a stepped care behavior therapy trial for TTM (Rogers et al., 2014). Step 1 consisted of web-based self-help and resulted in a small, statistically significant decrease in interviewer-rated symptoms but on the whole had weak effects. Step 2 was in-person HRT; uncontrolled data suggested Step 2 was on an average effective, and the stepped care program as a whole was highly acceptable to patients.

In the current study, we hypothesized that patients who achieved abstinence during treatment would be more likely to maintain their gains. Second, lower symptom severity at pre-treatment and at the time of initial treatment response were both hypothesized to predict better maintenance of gains. Third, we had no directional prediction for tests of whether patients who previously had at least one period of abstinence (and then relapse, by definition) prior to entering the study would be more likely to maintain gains. Fourth, it was hypothesized that abstinent patients who ceased experiencing urges to pull would be more likely to maintain gains than abstinent patients with residual urges to pull. Fifth, we hypothesized that higher intrinsic motivation for treatment at baseline and at the time of initial response to treatment would predict better maintenance. Finally, it was hypothesized that patients with greater treatment compliance would better maintain gains.

2. Methods

2.1. Design overview

The American University Institutional Review Board approved all recruitment and study procedures. After providing informed consent and completing baseline assessment, eligible participants were randomized to the immediate treatment ($n=30$) and waitlist conditions ($n=30$). In the immediate treatment condition, participants began the step 1 intervention, which consisted of 10 weeks of free access to StopPulling.com, self-help behavior therapy via an interactive website. In the waitlist condition, participants remained on a waitlist for 10 weeks before proceeding to step 1. Step 1 was followed by an in-person assessment (post-step 1), and at this time, participants chose whether to enter step 2 treatment (eight weeks of in-person habit reversal training) or to receive no further treatment. After eight weeks, all participants had a post-step 2 assessment, and then a 3-month follow-up assessment after a maintenance phase of no treatment for any participants. Participants could have initially met treatment response criteria at either the post-step 1 or post-step 2 assessment, at which point that individual participant's maintenance period began. Therefore, some participants could have received active treatment during their maintenance period. Data from all assessments conducted after the one at which the participant first met treatment response criteria were used to determine whether or not they had relapsed (see Fig. 1).

2.2. Participants

Sixty adults (57 female, 3 male) with TTM were enrolled in the study. The majority were Caucasian (75%), while 17% identified as African-American, 3% as Asian, 2% as Native Hawaiian/other Pacific Islander, and 3% as "other" race. One participant (2%) identified as Hispanic. The sample was highly educated, with 82% having completed college and 37% having completed graduate school. See Table 2 for additional demographic and clinical characteristics of the participants. Participants were recruited through newspaper advertisements, websites, and clinician referrals, and then screened by phone, followed by in-person evaluations with masters or doctoral level students. To qualify for the study, participants needed to 1) be at least 18 years old, 2) have regular internet access, and 3) meet DSM-IV criteria for TTM, except criteria B (increasing tension before pulling or attempting to resist) and C (pleasure, gratification, or relief when pulling) were not required (American Psychiatric Association, 2000). The B and C criteria were not required for this study because findings have not supported the incremental validity of these criteria as part of TTM diagnoses (Conelea et al., 2012). To augment external validity, we used the same exclusion criteria as are in effect for non-study users of StopPulling.com, the step 1 intervention, which are current suicidality, major depression, psychosis, severe anxiety, or substance abuse. In addition, participants were excluded if they were in concurrent psychotherapy focused on TTM or if they were taking a dosage of psychotropic medication for TTM that had not been stable for at least four weeks prior to study enrollment.

Analyses were conducted on participants who had completed at a minimum, the post-step 1 assessment after having received the step 1 intervention, and at least one assessment afterwards (either at post-step 2 or three-month follow-up), to measure the maintenance of gains. These criteria excluded 12% ($n=7$) of participants from the analyses.

2.3. Measures

Clinician and self-rated measures were administered at baseline, post-step 1, post-step 2, and 3-month follow-up (see Table 3 for assessment schedule). Interviewers were masters and doctoral level graduate students trained by the principal investigator (David A. F. Haaga). All assessments were videotaped, and a random sample of the interviews (20%) was rated by a second coder, who was masked as to the time point of the assessment and treatment condition.

2.3.1. TTM symptom history, severity, and impairment

Massachusetts General Hospital Hairpulling Scale (MGH-HPS; Keuthen et al., 1995). The MGH-HPS is a self-report instrument for the assessment of hair pulling severity during the preceding week. It consists of seven items including measures

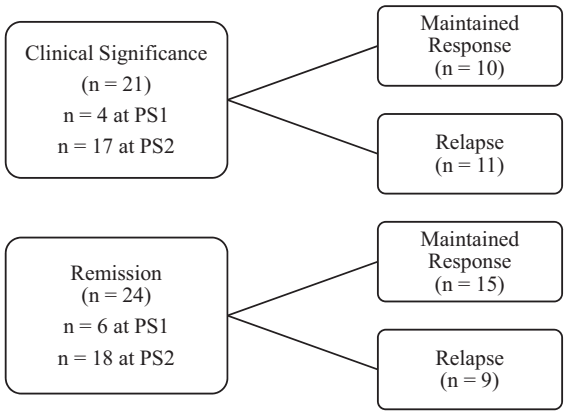

Fig. 1. Different outcomes following treatment response. The subgroups within Clinical Significance and Remission refer to the number of participants who met these treatment response criteria at the time of the post-step 1 or post-step 2 assessment.

Table 2
Demographic and clinical characteristics at baseline ($n=60$).

Characteristic	
Age	33.2 (10.9)
Age of TTM onset	11.5 (4.7)
Symptom duration (years)	21.2
Comorbid diagnoses (%)	
specific phobia	3
TTM severity (MGH-HPS)	16.9 (3.7)
Anxiety symptoms (BAI)	4.1 (4.6)
Depressive symptoms (BDI-II)	5.5 (5.1)
Prior experience with HRT (%)	17
Prior exposure to StopPulling.com (%)	3

Note: MGH-HPS=Massachusetts General Hospital Hairpulling Scale, BAI=Beck Anxiety Inventory, BDI=Beck Depression Inventory, HRT=Habit reversal training. Standard deviations appear in parentheses next to means.

Table 3
Assessment schedule.

Measure	Baseline	Post-step 1	Post-step 2	3-month follow-up
MGH-HPS	x	x	x	x
PITS	x	x	x	x
TDI-R	x	x	x	x
Trichotillomania course & treatment interview	x	–	–	–
SCID-I	x	–	–	–
BAI	x	x	x	x
BDI-II	x	x	x	x
CMOTS	x	x	x	–
Treatment utilization interview	–	x	x	x

Note: MGH-HPS=Massachusetts General Hospital Hairpulling Scale, PITS=The Psychiatric Institute Trichotillomania Scale, TDI-R=Trichotillomania Diagnostic Interview-Revised, SCID=Structured Clinical Interview for DSM-IV-TR Axis I Disorders, BAI=Beck Anxiety Inventory, BDI=Beck Depression Inventory, CMOTS=Client Motivation for Therapy Scale.

of frequency of hair pulling, resistance to and control over hair pulling, and distress. Items are rated on a severity scale ranging from 0 to 4, with overall severity scores ranging from 0 to 28. It has shown strong internal consistency (.74 for our sample at baseline) and test–retest reliability, with acceptable convergent and discriminant reliability, and sensitivity to change during treatment (Keuthen et al., 1995; O'Sullivan et al., 1995).

The Psychiatric Institute Trichotillomania Scale (PITS; Winchel et al., 1992). The PITS is a six-item, semi-structured interview that assesses TTM symptom severity. Items are each rated from 0 to 7, with higher scores reflective of greater severity (the total can range from 0 to 42). It has been found to have low internal consistency yet strong convergent validity with both self-report and other interviewer-rated TTM measures (Diefenbach, Tolin, Crocetto, Maltby, & Hannan, 2005). Our sample had low internal consistency on the PITS at baseline as well (alpha =.37). Twenty percent of the PITS interviews in this study were randomly selected for coding by a second rater, masked to treatment condition and assessment point. Item 6 could not be coded from video recordings. Thus, the sum of items 1–5 were evaluated for reliability and found to have a high correlation between interviewer and videocoder (r =.95).

Trichotillomania Diagnostic Interview-Revised (TDI-R; Rothbaum & Ninan, 1994). The TDI-R is a clinician-based, semi-structured interview modeled after the SCID, consisting of 3-point ratings of responses to items assessing the DSM-IV criteria for TTM. The original TDI was based on DSM-III-R criteria and subsequently revised to assess DSM-IV criteria by Keuthen et al. (2010). Participants needed to meet the DSM-IV criteria for TTM as assessed in this measure in this study at baseline, except criteria B and C were not required. Among the random sample (20%) of these interviews that was coded by a second rater, overall agreement was 92%, kappa =.77.

Trichotillomania Course and Treatment Interview (Haaga et al., *unpublished measure).* This structured interview was created for this study and consists of questions about the participant's course of TTM symptoms, use of medication, in-person therapy, and online self-help. It was used in this report for information concerning past episodes of abstinence from hair pulling.

2.3.2. Comorbid symptoms

Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition with Psychotic Screen (SCID-I/P with Psychotic Screen; First, Spitzer, Gibbon & Williams, 2002). The SCID-I/P is a semi-structured interview which was used to diagnose comorbid Axis I disorders in our sample. Among the random sample (20%) of these interviews that was coded by a second rater, 100% of the diagnoses made were agreed upon. *Beck Anxiety Inventory* (BAI; Beck, Epstein, Brown, & Steer, 1988). This 21-item self-report measure was used to assess comorbid anxiety symptoms. The BAI has shown high internal consistency, test–retest reliability, and construct validity.

Beck Depression Inventory, Second Edition (BDI-II; Beck, Steer, & Brown, 1996). The BDI-II is a 21-item self-report scale of the severity of depressive symptoms over the course of the preceding two weeks. It has been shown to have strong internal consistency and high test–retest reliability, and the number of symptoms endorsed by patients correlates with the number endorsed in clinical interviews (Sprinkle et al., 2002).

2.3.3. Intrinsic motivation for therapy

Client Motivation for Therapy Scale (CMOTS; Pelletier, Tuson, and Haddad, 1997). The CMOTS is a 24-item self-report questionnaire that measures several types of motivation, based on the conceptualization of motivation by Deci and Ryan, (1985). The intrinsic motivation subscale was used in this study, and is indexed by a 4-item subscale in which scores range from 4 to 28, with higher scores indicating greater intrinsic motivation. This measure has been shown to correlate positively with one's intent to remain in therapy and perception of therapy as important (Pelletier et al., 1997).

2.3.4. Treatment utilization and compliance

Treatment Utilization Interview. The treatment utilization section of the *Trichotillomania Course and Treatment History Interview* was administered at the post-step 1, post-step 2, and follow-up assessments.

Adherence to StopPulling.com. Adherence to the StopPulling.com program was measured by objectively tracking the number of days (0–70) on which participants logged on to the site and entered data during their period of free access to StopPulling.com.

HRT homework compliance. Homework assignments were rated for completion by the participant's therapist as 0 ("Not done"), 1 ("Reportedly done but not documented"), 2 ("Partially completed and documented"), or 3 ("Fully or almost fully completed and documented").

2.3.5. Criteria of abstinence from hair pulling

Periods of abstinence prior to treatment and post-treatment were both measured. Abstinence periods experienced prior to treatment were operationalized as any period of at least two weeks during which the participant reports no hair was pulled, as measured by the Trichotillomania Course and Treatment Interview at the baseline assessment. Current periods of abstinence (during post-treatment assessments) were measured using the same criterion that was used in Keijsers et al. (2006), a score of "0" on item 4 of the MGH-HPS (frequency of hair pulling), "this week I did not pull my hair". In a secondary analysis, current periods of abstinence were measured using item 1 of the TDI-R, "do you pull out hair anywhere on your body other than for cosmetic reasons?"

2.3.6. Criteria for residual urges

Participants needed to meet two criteria to be considered as having residual urges. In addition to the criteria for abstinence being satisfied, participants must have endorsed a score of 1 or above (1=occasional urge, 2=urge to pull often, 3=very often, 4=near constant) on item 1 of the MGH-HPS (frequency of urges).

2.3.7. Criteria for treatment response

Several operational definitions were used in this study to measure different outcomes following treatment response (see Fig. 1). *Response* was defined as, *improvement of symptoms with treatment* and was measured by two methods: *clinical significance* or *remission. Clinical significance* was measured according to guidelines by Jacobson and Truax (1991), which specify that participants must meet both (a) recovery of normal-range functioning, which was assessed with a post-treatment cut-off score of 9 or below on the MGH-HPS in our sample as it was calculated as more than 2 standard deviations below our sample's mean MGH-HPS scores at baseline, and (b) reliable change (RC), in which the RC index was $\geq$1.96. Reliable change was calculated to be a change of 6 or more points in the direction of improvement on the MGH-HPS. As recommended by Lambert, Hansen, and Bauer (2008), we used internal consistency (.74 in our baseline data) of the MGH-HPS as the reliability value needed to calculate the RC index. The MGH-HPS was chosen as the primary outcome measure in these calculations because of its widespread use and validation (e.g., Diefenbach et al., 2006; van Minnen et al., 2003). Forty percent

($n=21$) of the participants in the study met clinical significance criteria at least once.

Remission was defined as *improvement of symptoms such that the individual no longer meets criteria for TTM and has minimal symptoms at most*. This was measured by the lack of meeting the TTM diagnostic criteria for the study as measured by the TDI-R, which does not require a minimum time frame for not meeting diagnosis. This was fulfilled by 45% ($n=24$) of participants at least once.

The two measures of treatment response (*clinical significance*, measured by the MGH-HPS and *remission*, measured by the TDI) overlapped significantly but not completely. Of those with complete data on both, 70% of the participants who met either CS or remission criteria met both (kappa =.58).[1]

2.3.8. Criteria for maintenance of response

Participants who continued meeting clinical significance criteria were considered as having *maintained response*, yet if they ceased meeting these criteria, they experienced *relapse*. Fifty-two percent ($n=11$) of the participants who met *clinical significance* criteria ($n=21$) relapsed. Participants who continued not meeting TTM diagnostic criteria had *maintained response*, whereas those who began meeting TTM diagnostic criteria again had experienced *relapse*, which was defined as the return of symptoms during a period of remission to the extent that the TTM diagnostic criteria are met. Thirty-eight percent ($n=9$) of the participants who had remitted ($n=24$) relapsed during the study.

2.4. Characterization of significant predictor data

Probability of Treatment Benefit (PTB) charts are presented to characterize the implications of significant predictors in this study, using the method devised by Lindhiem et al. (2012). These charts quantify the probability of maintaining gains from the treatment as a function of participants' scores on the predictors. The original method used to create this chart was slightly revised to predict the probability of maintenance of gains over time given some initial response to treatment, rather than the prediction of this initial response. Participants were stratified according to what level of the predictor variables they endorsed, with continuous variables dichotomized (e.g., intrinsic motivation), and the probabilities of the participants in each of these strata meeting the outcome variables (e.g., maintained response) were computed. Logistic models were used to predict these probabilities and their corresponding 95% confidence intervals. To illustrate the utility of depicting prognostic effects via PTB charts, a study reporting only that conscientiousness predicts benefit from exposure and response prevention for OCD with a particular effect size might be very difficult for the individual patient or clinician to interpret. If this same relation was recast in a PTB chart showing that, for instance, "adults diagnosed with OCD and scoring 35–39 on Measure X for conscientiousness have a 65% chance (95% confidence interval 50–80) of gaining clinically significant benefit from ERP, compared to just 35% chance (95% CI 21–49) for those scoring 25–29 on Measure X," the treating clinician and the patient could more readily use the information in deciding how to proceed. The application of PTB charts in our study entailed quantifying the likelihood of participants maintaining their gains, as opposed to initial treatment response.

2.5. Procedure

2.5.1. Step 1: Self-help behavior therapy

While in step 1 of the study, all participants ($n=60$) had 10 weeks of free access to behavior therapy from StopPulling.com. The website takes participants through several modules in which they are first asked to identify their hair-pulling triggers by self-monitoring the situations, feelings, behaviors, and thoughts which lead to pulling. As the modules progress, participants are asked for more specific information, such as post-pulling behavior. Subsequently, each week, three coping strategies that seem to be relevant to the client are presented to the participants, and participants are asked to self-monitor their use of these strategies, along with continuing to monitor their urges and pulling, while setting weekly goals and self-administering rewards after meeting these goals.

At the end of step 1, all participants were offered step 2, eight weekly sessions of in-person behavior therapy (specifically, Habit Reversal Training). Regardless of whether or not participants chose to receive the HRT, they still continued to participate in subsequent assessments.

2.5.2. Step 2: Habit Reversal Training (HRT)

Step 2 consisted of eight sessions of individual in-person behavior therapy with a doctoral student in a university outpatient clinic. Seventy-six percent ($n=41$) of the 54 participants who completed the post-step 1 assessment chose to receive the step 2 intervention, and they attended a mean of 7.61 of eight scheduled sessions. There were seven therapists, who ranged from being in the first to the fifth year of their clinical training and had been trained and supervised by the principal investigator of the study. All HRT sessions were videotaped and used for adherence assessment and in supervision. The HRT was based on a modified version of the manual written by Stanley and Mouton (1996). This manual was modified to be for individual therapy, extend the length of treatment, and with an increased emphasis on stimulus control. This protocol focused on self-monitoring, awareness training, stimulus control, and stimulus-response and competing response interventions. A 10% random sample of the sessions was reviewed for therapist adherence to the protocol by two raters who were not the therapists for the sessions to be rated, but who were made familiar with the manual for HRT (97% overall agreement, kappa =.78).

2.6. Statistical analyses

The Statistical Package for the Social Sciences (SPSS, version 19) was used for all statistical analyses. Fisher's exact test was used to compare categorical variables for the hypotheses regarding abstinence and residual urges. Abstinence was also evaluated as a predictor of hair pulling severity at follow-up by the independent sample *t*-test. Logistic regressions were performed to determine whether TTM symptom severity and intrinsic motivation predicted maintenance. Logistic models were used to calculate predicted probabilities and their corresponding confidence intervals for the PTB chart. All tests of significance were two-tailed. Nominal, per-comparison, p values are reported, but for our primary analyses of predictors we used a Bonferroni correction in order to control familywise Type I error rates. In particular, we considered $p < .025$ (.05/2) results significant, as each predictor was tested twice (in the subsample who achieved clinically significant response and in the subsample who remitted). Prior to conducting the analyses, it was confirmed that all statistical assumptions were met.

3. Results

3.1. Predictors of maintenance

3.1.1. Post-treatment abstinence

Post-treatment abstinence from hair pulling was a strong and significant (even with Bonferroni correction) predictor of maintenance among the subsample who met clinical significance criteria. Specifically 77% of abstainers maintained their gains, whereas 0% of those who showed a clinically significant response but not complete abstinence maintained clinical significance ($p=.003$, Fisher's exact test). However, among those who remitted, post-treatment abstinence did not predict maintenance: 73% of those who abstained maintained remission, while 54% of those who did not abstain maintained remission. ($p=.423$, Fisher's exact test; Table 4).

Post-treatment abstinence (predictor) and maintenance of clinically significant response (outcome) are both indexed via the self-report MGH-HPS. Thus, method variance may contribute to the significant predictive effect, particularly in light of our not finding abstinence to predict maintenance of remission, which is operationalized via structured interview (TDI-R). Accordingly, a secondary analysis was conducted, replacing MGH-HPS with the TDI-R as the abstinence measure. This secondary analysis with the TDI-R yielded a non-significant result in which 75% of those abstinent maintained their gains, whereas 33% of treatment responders who were not abstinent post-treatment maintained their gains ($p=.085$, Fisher's exact test).

A secondary analysis of the hypothesis that post-treatment abstinence predicts maintenance was conducted to replicate the methods of Keijsers et al. (2006). The sample was divided into participants who were abstainers at either the post-step 1 or post-step 2 assessment ($n=14$) and non-abstainers at either post-treatment assessment ($n=36$). MGH-HPS scores of the abstainers at three-month follow-up ($M=7.9$, $SD=6.0$) were significantly lower ($t(48)=5.55$, $p<.001$; $d=1.60$) than the follow-up scores of those who had been non-abstainers ($M=16.3$, $SD=4.3$) at the conclusion of treatment. These results replicated those found by Keijsers et al. (2006).

3.1.2. Pre-treatment TTM severity

Logistic regression analyses were conducted using pre-treatment TTM severity scores as a predictor and whether or not participants

[1] Due to missing data, 23 of 24 remitters had data relevant to clinical significance criteria; 19 of those meeting clinical significance had remission criteria data.

maintained their gains served as the dependent variable. Pre-treatment TTM severity did not predict maintenance among participants who met clinical significance criteria during treatment (odds ratio (OR),.84; 95% confidence interval (CI):.69, 1.19; $p = .326$) or among those who remitted during the course of treatment (OR, .99; 95% CI:.78, 1.26; $p = .914$).

3.1.3. TTM severity at time of initial response

TTM severity at the time of initial response to treatment had a significant (Bonferroni-adjusted) inverse relationship with maintenance among participants who had met clinical significance criteria during the course of treatment (OR, .41, 95% CI:.20,.83; $p = .013$). Participants who maintained clinically significant response had lower MGH-HPS scores at the time they initially met clinical significance criteria ($M = 3.90$, $SD = 1.91$) than participants who did not maintain clinical significance criteria ($M = 7.25$, $SD = 1.39$). Among participants who had remitted, TTM severity at the time of initial response did not predict maintenance (OR, 1.10, 95% CI:.93, 1.29; $p = .270$).

Pre-treatment symptom severity and maintenance of clinically significant response were both measured using the MGH-HPS. Consequently, method variance could have been a factor in the significant predictive effect. To address this, a secondary analysis was performed using the PITS as an alternative to the MGH-HPS TTM severity as measured by the PITS at the time of initial clinically significant response which did not predict maintenance of clinically significant response (OR, .90, 95% CI:.75, 1.09; $p = .284$). The PITS scores at the time of initial clinically significant response among who maintained clinically significant response ($M = 10.70$, $SD = 5.91$) tended to be lower than those who did not maintain their gains ($M = 13.38$, $SD = 4.37$).

3.1.4. History of abstinence and relapse prior to treatment

Participants ($N = 60$) reported a wide range of the longest abstinence periods that they had experienced in two weeks or more prior to treatment. 60% ($n = 36$) of the sample reported having abstained for at least two weeks during their lives prior to entering the study, with the median longest period of abstinence being 98 days (range, 14–3650 days). Among those who showed clinically significant response, having a prior period of abstinence did not predict maintaining vs. failing to maintain that response ($p = 1.00$, Fisher's exact test). Likewise, among the remitted subsample, prior abstinence did not predict maintenance ($p = .669$, Fisher's exact test; Table 5).

3.1.5. Residual urges

Residual urges were not found to predict relapse in either participants who met clinical significance criteria ($p = .528$, Fisher's exact test) or participants who remitted ($p = .491$, Fisher's exact test; Table 6).

3.1.6. Pre-treatment intrinsic motivation

Higher intrinsic motivation for treatment at baseline was not found to predict better maintenance of gains at follow-up among participants meeting clinical significance criteria (OR, 1.00, 95% CI:.85, 1.18; $p = 1.00$) or participants who remitted (OR, 1.20, 95% CI:.98, 1.48; $p = .076$).

3.1.7. Intrinsic motivation at time of initial response

Higher intrinsic motivation for treatment at the time of initial response was also not found to predict the maintenance of gains among participants meeting clinical significance criteria (OR, .97; 95% CI:.81, 1.15; $p = .698$) and those who remitted (OR, 1.07; 95% CI:.93, 1.23; $p = .338$).

3.1.8. Treatment compliance

The mean number of days logged on StopPulling.com did not predict maintenance of gains for participants who met clinical significance criteria (OR: 1.0, 95% CI:.98, 1.0, $p = .568$) or remitters (OR: 1.0, 95% CI:.97, 10, $p = .884$). Similarly, the mean adherence scores for HRT homework compliance did not predict maintenance of gains for those who met clinical significance (OR:.43, 95% CI:.08, 2.2, $p = .312$) or remission (OR:.44, 95% CI: −.07, 2.8, $p = .384$).

3.2. Probability of treatment benefit charts

Table 7 displays the implications of the relations of the significant predictors in a PTB chart. Thus, for instance, a clinician could infer that a TTM patient scoring 0 to 4 on the MGH-HPS at the time of achieving clinically significant response would have an 88% probability of maintaining that response through 3-month follow-up, compared to just 23% chance for a patient scoring 5 to 9 on the MGH-HPS at the time of clinically significant response. Note that the regression analysis that corresponds with this result in the PTB chart was calculated using the MGH-HPS as a continuous variable, whereas the result in the PTB chart was calculated by dichotomizing the MGH-HPS into two levels (0 to 4 and 5 to 9) to illustrate the implications of the finding. In addition, by using this chart a clinician could infer that a patient who is abstinent at the conclusion of treatment has 77% chance of maintaining gains, but 0% chance if not abstinent at the conclusion of treatment. This chart shows the fairly wide confidence intervals surrounding these estimates, which underscores the need for large-sample research to increase the precision of such estimates.

4. Discussion

The current study investigated predictors of maintenance vs. relapse among responders to behavior therapy for TTM. Lower TTM symptom severity at the time of initial response and abstinence from hair pulling at the conclusion of behavior therapy predicted increased probability of maintenance. Pre-treatment TTM severity, abstinence periods prior to treatment, residual urges after having achieved abstinence, intrinsic motivation before receiving treatment and during initial response to treatment, and treatment compliance were not significant predictors of maintenance. Replication in larger, more statistically powerful studies is needed in order to establish definitively whether or not these potential predictors are beneficial or detrimental to patients in treatment for TTM.

Abstinence at the conclusion of treatment predicted maintenance among those who made clinically significant improvement. This replicates the findings of both Lerner et al. (1998) and Keijsers et al. (2006), in which follow-up hair pulling severity was lower among post-treatment abstainers than among non-abstainers These findings are similar to those obtained in treatment studies of other disorders as well, such as a study of cognitive behavior therapy for depression which showed that patients who were fully

Table 4
Post-treatment abstinence as a predictor of maintenance vs. relapse

	Clinical significance		Remission	
	Post-tx abstinence	No post-tx abstinence	Post-tx abstinence	No post-tx abstinence
Maintained	10	0	8	7
Relapsed	3	7	3	6
Maintenance rate (%)	77	0	73	54

Note: Post-tx = post-treatment.

recovered were at a lower risk for relapsing than patients who had only partially recovered (Thase et al. 1992). Future research testing alternative explanations of this finding would be very useful. Conceptual and empirical analyses of predictors of return of fear after exposure therapy for anxiety disorders have incorporated, and benefitted from, basic learning research (e.g., Abramowitz, 2013). Limitations of this analysis included method variance (abstinence and clinical significance criteria originated from the same measure) and inconsistent results such that these findings were not replicated among remitters, only those who had shown clinically significant improvement.

TTM symptom severity at baseline did not significantly predict better gain maintenance.

Lower TTM symptom severity at the time of initial response predicted maintenance among those who experienced clinically significant improvement, but not among those who remitted. To our knowledge, this is the first study to examine the predictive value of TTM severity at the time of initial response in maintaining gains. It elaborates upon the findings of Lerner et al. (1998) that higher levels of pre-treatment TTM severity predicted higher symptom severity in long-term. A limitation of the methods used for this analysis was that the data for both TTM symptom severity and clinical significance criteria originated from the same measure.

Having had at least one period of abstinence and then relapse prior to entering treatment did not predict maintenance. In the sample, 50% of participants who had prior abstinence and achieved clinically significant improvement sustained that improvement, compared to 40% without prior abstinence (Table 5). Based on these results, it appears that achieving abstinence prior to therapy may not have an effect on one's probability of maintaining gains. It could be the case that prior abstinence does not lead to useful knowledge for maintaining improvement in the future. Alternatively, prior periods of abstinence may be useful learning experiences with respect to how to maintain gains, but this effect may be offset by a selection effect such that hair pullers who are still seeking treatment after prior abstinence periods are people who are especially prone to relapsing. Regardless of the best explanation, our results are inconsistent with Schachter's (1982) hypothesis that incremental learning would make prior relapses a positive predictor of maintenance of change in the current treatment episode.

Many patients who achieve abstinence at the conclusion of treatment continue to struggle with residual urges to pull. In the current study, 71% of the participants who were abstinent at the conclusion of treatment ($n=14$) had residual urges to pull. These participants were not found to be more likely to relapse than participants without residual urges. This is at odds with prior research done in smoking cessation trials suggesting that residual urges lead to a quicker relapse (e.g., Doherty et al., 1995) and theories about the likelihood of urges leading to relapse (e.g. Muraven & Baumeister, 2000). However, it seems plausible that with a larger sample size, the results would have differed. Among participants who had achieved clinically significant improvement or remitted after receiving treatment, 100% of participants without residual urges maintained these gains, whereas only 63–70% of those who had urges maintained their gains (Table 6).

Table 5
Pre-treatment abstinence as a predictor of maintenance vs. relapse

	Clinical significance		Remission	
	Prior abstinence	No prior abstinence	Prior abstinence	No prior abstinence
Maintained	8	2	10	5
Relapsed	8	3	7	2
Maintenance rate (%)	50	40	59	71

Table 6
Residual urges as a predictor of maintenance vs. relapse.

	Clinical significance		Remission	
	Urges	No urges	Urges	No urges
Maintained	7	3	5	3
Relapsed	3	0	3	0
Maintenance rate (%)	70	100	63	100

Table 7
Probability of treatment benefit chart for TTM severity during initial response.

	Probability of treatment benefit (Maintenance of gains)
	Clinical significance (95% CI)
MGH-HPS: 0–4	88% (46–98%)
MGH-HPS: 5–9	23% (8–52%)
Abstinent	77% (48–92%)
Not abstinent	0% (0–100%)

Note: Post-treatment abstinence was measured by a score of "0" on item 4 of the MGH-HPS at the conclusion of treatment. MGH-HPS scores were measured at the time of meeting clinically significant improvement. The probability of treatment benefit is the percent of participants who made clinically significant improvement and maintained this improvement.

The hypotheses that higher intrinsic motivation for treatment at baseline and at the time of initial response to treatment would predict better maintenance were not supported. This lack of significant findings is in contrast to previous findings that showed intrinsic motivation correlated positively with maintenance of changes in health behaviors (e.g., Ryan & Deci, 2000). Our sample appeared to have adequate intrinsic motivation for treatment, with mean total scores on the CMOTS intrinsic motivation subscale of 16.9 ($SD=5.3$). By comparison, a sample of participants with generalized anxiety disorder that was randomized into different treatment conditions had means of 17.06 ($SD=6.05$) through 19.84 ($SD=4.31$) among groups before treatment (Westra, Arkowitz, & Dozois, 2009). Because our sample appeared to be adequately motivated for treatment, it is possible that motivation was not a significant predictor since the CMOTS is generically related to one's motives to be in therapy, and does not measure one's actual state of readiness to reduce hair pulling. As an illustration of this, a sample item from the CMOTS is, "[I am interested in being in therapy] for the satisfaction I think I will have when I try to achieve my personal goals in the course of therapy."

Contrary to our expectations that increased frequency of skills practice would result in better maintenance, neither greater adherence to StopPulling.com nor greater completion of homework assignments in HRT predicted maintenance in our sample. One possible explanation for this is that frequency of skills practice is not as important as the quality of the skills practice, as was shown by Schmidt and Woolaway-Bickel (2000) in a CBT trial for panic disorder.

4.1. Limitations and future research

Several limitations should be considered in interpreting the results of this study. First, as noted earlier, method variance could have been a factor in the findings involving the MGH-HPS as a measure of both the predictor and outcome variables. Second, our subsample of treatment responders was small, detracting from the statistical power of analyses predicting maintenance vs. relapse. Third, the follow-up period of three months was relatively short.

Fourth, the sample was highly educated (82% finished college). Future studies of predictors of relapse after TTM treatment would benefit from use of larger and more diverse samples, multimodal measurement, and longer follow-up periods.

5. Conclusion

Of the possible predictors evaluated, abstinence at the conclusion of treatment and lower hair pulling severity at the time of initial response were found to significantly predict maintenance. Replications of this research are needed to determine the usefulness of these possible predictors in identifying relapse-prone patients, with the ultimate aim of improving clinical decision-making and developing strategies to help these patients better maintain treatment gains. Probability of Treatment Benefit charts should be used more widely in TTM research to clarify the practical prognostic implications of studies, such as those in this study.

Another important result from this study is that the analyses yielded different findings about the same research question depending on which definition of treatment response was used. For example, post-treatment abstinence significantly predicted maintenance only among those who had showed clinically significant change, but not among those who had remitted. Table 1 presents the variety of operational definitions used by TTM researchers for the concept of "treatment response" across studies. It is important that gold standard outcome variables are established in TTM treatment studies to afford future cross-study comparisons. Nelson et al., 2014 evaluated the relative validity of seven indicators of treatment response (complete abstinence, and as measured by the MGH-HPS and PITS $>=25\%$ symptom reduction, recovery of normal functioning, and clinical significance) with the full sample in this study, and found that clinically significant improvement on the MGH-HPS had the strongest results. They also recommended the use of data on complete abstinence as a supplementary metric.

Role of funding sources

This research was supported by National Institute of Mental Health Grant 1R15MH086852-01. NIMH had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication.

Contributors

Authors A and D designed the study and wrote the first draft of the manuscript. Author A also conducted literature searches. Authors A and C conducted the statistical analysis. Authors B and D designed and collected data for the original NIMH-funded clinical trial. All authors contributed to and have approved the final manuscript.

Conflict of Interest

All authors declare that they have no conflicts of interest.

Acknowledgments

The authors wish to thank Kate Stewart, who assisted in the proof-reading of the manuscript, and Charley Mansueto, who consulted on the stepped care project as a whole.

References

Abramowitz, J. S. (2013). The practice of exposure therapy: relevance of cognitive-behavioral theory and extinction theory. *Behavior Therapy*, *44*, 548–558. http://dx.doi.org/10.1016/j.beth.2013.03.003.

American Psychiatric Association (2000). (text revision). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.

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

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory Manual* (2nd ed.). San Antonio, TX: The Psychological Corporation.

Bloch, M. H., Landeros-Weisenberger, A., Dombrowski, P., Kelmendi, B., Wegner, R., Nudel, J., & Coric, V. (2007). Systematic review: pharmacological and behavioral treatment for trichotillomania. *Biological Psychiatry*, *62*(8), 839–846. http://dx.doi.org/10.1016/j.biopsych.2007.05.019.

Brewer, J. A., & Potenza, M. N. (2008). The neurobiology and genetics of impulse control disorders: relationships to drug addictions. *Biochemical Pharmacology*, *75*(1), 63–75. http://dx.doi.org/10.1016/j.bcp.2007.06.043.

Carroll, K. M., & Rousnaville, B. J. (2006). Behavioral therapies: the glass would be half full if only we had a glass. In: W. R. Miller, & K. M. Carroll (Eds.), *Rethinking substance abuse: What the science shows, and what we should do about it* (pp. 223–239). New York: Guilford Press.

Conelea, C. A., Walther, M. R., Flessner, C. A., Woods, D. W., Franklin, M. E., Keuthen, N. J., & Piacentini, J. C. (2012). The incremental validity of criteria B and C for diagnosis of trichotillomania in children and adults. *Journal of Obsessive-Compulsive and Related Disorders*, *1*(2), 98–103. http://dx.doi.org/10.1016/j.jocrd.2012.01.004.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.

Diefenbach, G. J., Mouton-Odum, S., & Stanley, M. A. (2002). Affective correlates of trichotillomania. *Behaviour Research and Therapy*, *40*(11), 1305–1315. http://dx.doi.org/10.1016/S0005-7967(02)00006-2.

Diefenbach, G. J., Reitman, D., & Williamson, D. A. (2000). Trichotillomania: a challenge to research and practice. *Clinical Psychology Review*, *20*, 289–309. http://dx.doi.org/10.1016/S0272-7358(98)00083-X.

Diefenbach, G. J., Tolin, D. F., Crocetto, J., Maltby, N., & Hannan, S. (2005). Assessment of trichotillomania: a psychometric evaluation of hair-pulling scales. *Journal of Psychopathology and Behavioral Assessment*, *27*, 169–178. http://dx.doi.org/10.1007/s10862-005-0633-7.

Diefenbach, G. J., Tolin, D. F., Hannan, S., Maltby, N., & Crocetto, J. (2006). Group treatment for trichotillomania: behavior therapy versus supportive therapy. *Behavior Therapy*, *37*(4), 353–363. http://dx.doi.org/10.1016/j.beth.2006.01.006.

Doherty, K., Kinnunen, T., Militello, F. S., & Garvey, A. J. (1995). Urges to smoke during the first month of abstinence: relationship to relapse and predictors. *Psychopharmacology*, *119*, 171–178. http://dx.doi.org/10.1007/BF02246158.

Edelman, R. E., & Chambless, D. L. (1995). Adherence during sessions and homework in cognitive-behavioral group treatment of social phobia. *Behaviour Research and Therapy*, *33*, 573–577.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured clinical interview for DSM-IV-TR axis I disorders, (Research Version, Patient Edition with Psychotic Screen)*. New York: Biometrics Research, New York State Psychiatric Institute.

Grant, J. E., Odlaug, B. L., & Potenza, M. N. (2007). Addicted to hair pulling? How an alternate model of trichotillomania may improve treatment outcome. *Harvard Review of Psychiatry*, *15*(2), 80–85. http://dx.doi.org/10.1080/10673220701298407.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19. http://dx.doi.org/10.1037//0022-006X.59.1.12.

Keijsers, G. P. J., van Minnen, A., Hoogduin, C. A. L., Klaassen, B. N. W., Hendriks, M. J., & Tanis-Jacobs, J. (2006). Behavioural treatment of trichotillomania: two-year follow-up results. *Behaviour Research and Therapy*, *44*(3), 359–370. http://dx.doi.org/10.1016/j.brat.2005.03.004.

Keuthen, N. J., O'Sullivan, R. L., Ricciardi, J. N., Shera, D., Savage, C. R., Borgman, A. S., & Baer, L. (1995). The Massachusetts General Hospital (MGH) Hairpulling Scale: I. development and factor analysis. *Psychotherapy and Psychosomatics*, *64*, 141–145. http://dx.doi.org/10.1159/000289003.

Keuthen, N. J., Rothbaum, B. O., Falkenstein, M. J., Meunier, S., Timpano, K. R., Jenike, M. A., & Welch, S. S. (2011). DBT-enhanced habit reversal treatment for trichotillomania: 3-and 6-month follow-up results. *Depression and Anxiety*, *28*(4), 310–313. http://dx.doi.org/10.1002/da.20778.

Keuthen, N. J., Rothbaum, B. O., Fama, J., Altenburger, E., Falkenstein, M. J., Sprich, S. E., & Welch, S. S. (2012). DBT-enhanced cognitive-behavioral treatment for trichotillomania: a randomized controlled trial. *Journal of Behavioral Addictions*, *1*(3), 106–114. http://dx.doi.org/10.1556/JBA.1.2012.003.

Keuthen, N. J., Rothbaum, B. O., Welch, S. S., Taylor, C., Falkenstein, M., Heekin, M., & Jenike, M. A. (2010). Pilot trial of dialectical behavior therapy-enhanced habit reversal for trichotillomania. *Depression and Anxiety*, *27*(10), 953–959. http://dx.doi.org/10.1002/da.20732.

Lambert, M. J., Hansen, N. B., & Bauer, S. (2008). Assessing the clinical significance of outcome results. In: A. M. Nezu, & C. M. Nezu (Eds.), *Evidence-based outcome research: a practical guide to conducting randomized controlled trials for psychosocial interventions* (pp. 359–378). New York: Oxford University Press.

Lerner, J., Franklin, M. E., Meadows, E. A., Hembree, E., & Foa, E. B. (1998). Effectiveness of a cognitive behavioral treatment program for trichotillomania: an uncontrolled evaluation. *Behavior Therapy*, *29*, 157–171. http://dx.doi.org/10.1016/S0005-7894(98)80036-1.

Lindhiem, O., Kolko, D. J., & Cheng, Y. (2012). Predicting psychotherapy benefit: a probabilistic and individualized approach. *Behavior Therapy*, *43*, 381–392. http://dx.doi.org/10.1016/j.beth.2011.08.004.

Marlatt, G. A., & Witkiewitz, K. (2005). Relapse prevention for alcohol and drug problems. In: G. A. Marlatt, & D. M. Donovan (Eds.), *Relapse prevention: maintenance strategies in the treatment of addictive behaviors* (pp. 1–44). New York: Guilford Press.

Miller, W. R., & Carroll, K. M. (2006). Drawing the science together: ten principles, ten recommendations. In: W. R. Miller, & K. M. Carroll (Eds.), *Rethinking substance abuse: what the science shows, and what we should do about it* (pp. 293–311). New York: Guilford Press.

Mouton, S., & Stanley, M. (1996). Habit reversal training for trichotillomania: a group approach. *Cognitive and Behavioral Practice*, *3*(1), 159–182. http://dx.doi.org/10.1016/S1077-7229(96)80036-8.

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: does self-control resemble a muscle? *Psychological Bulletin*, *126*(2), 247–259. http://dx.doi.org/10.1037/0033-2909.126.2.247.

Nelson, S. O., Rogers, K. R., Rusch, N., McDonough, L., Malloy, E. J., Falkenstein, M. J., Banis, M., & Haaga, D. A. F. (2014). Validating indicators of treatment response: application to trichotillomania. *Psychological Assessment*, *26*(3), 857–864.

Ninan, P. T., Rothbaum, B. O., Marsteller, F. A., Knight, B. T., & Eccard, M. B. (2000). A placebo-controlled trial of cognitive-behavioral therapy and clomipramine in trichotillomania. *The Journal of Clinical Psychiatry*, *61*(1), 47–50.

O'Sullivan, R. L., Keuthen, N. J., Hayday, C. F., Ricciardi, J. N., Buttolph, M. L., Jenike, M. A., & Baer, L. (1995). The Massachusetts General Hospital (MGH) hairpulling scale: 2. Reliability and validity. *Psychotherapy and Psychosomatics*, *64*, 146–148. http://dx.doi.org/10.1159/000289004.

Partos, T. R., Borland, R., Yong, H. H., Hyland, A., & Cummings, K. M. (2013). The quitting rollercoaster: how recent quitting history affects future cessation outcomes (data from the International Tobacco Control 4-country cohort study). *Nicotine & Tobacco Research*, *15*, 1578–1587. http://dx.doi.org/10.1093/ntr/ntt025.

Pelletier, L. G., Tuson, K. M., & Haddad, N. K. (1997). Client motivation for therapy scale: a measure of intrinsic motivation, extrinsic motivation, and amotivation for therapy. *Journal of Personality Assessment*, *68*(2), 414–435. http://dx.doi.org/10.1207/s15327752jpa6802_11.

Rogers, K., Banis, M., Falkenstein, M. J., Malloy, E. J., McDonough, L., Nelson, S., Rusch, N., & Haaga, D. A. F. (2014). Stepped care in the treatment of trichotillomania. *Journal of Consulting and Clinical Psychology*, *82*, 361–367. http://dx.doi.org/10.1037/a0035744.

Rothbaum, B. O., & Ninan, P. T. (1994). Assessment of trichotillomania. *Behaviour Research and Therapy*, *32*(6), 651–662. http://dx.doi.org/10.1016/0005-7967(94)90022-1.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68–78. http://dx.doi.org/10.1037/0003-066X.55.1.68.

Schachter, S. (1982). Recidivism and self-cure of smoking and obesity. *American Psychologist*, *37*(4), 436–444. http://dx.doi.org/10.1037/0003-066X.37.4.436.

Schmidt, N. B., & Woolaway-Bickel, K. (2000). The effects of treatment compliance on outcome in cognitive-behavioral therapy for panic disorder: quality versus quantity. *Journal of Consulting and Clinical Psychology*, *68*, 13–18.

Sprinkle, S. D., Lurie, D., Insko, S. L., Atkinson, G., Jones, G. L., Logan, A. R., & Bissada, N. N. (2002). Criterion validity, severity cut scores, and test-retest reliability of the Beck Depression Inventory-II in a university counseling center sample. *Journal of Counseling Psychology*, *49*(3), 381.

Stanley, M. A., & Mouton, S. G. (1996). Trichotillomania treatment manual. In: M. Hersen, & V. VanHasselt (Eds.), *Sourcebook of psychological treatment manuals for adult disorders* (pp. 657–687). New York: Plenum Press.

Thase, M. E., Simons, A. D., McGeary, J., Cahalane, J. F., Hughes, C., Harden, T., & Friedman, E. (1992). Relapse after cognitive behavior therapy of depression: potential implications for longer courses of treatment. *The American Journal of Psychiatry*, *149*(8), 1046–1052.

Twohig, M. P., & Woods, D. W. (2004). A preliminary investigation of acceptance and commitment therapy and habit reversal as a treatment for trichotillomania. *Behavior Therapy*, *35*(4), 803–820. http://dx.doi.org/10.1016/S0005-7894(04)80021-2.

van Minnen, A., Hoogduin, K. A. L., Keijsers, G. P. J., Hellenbrand, I., & Hendriks, G. J. (2003). Treatment of trichotillomania with behavioral therapy or fluoxetine: a randomized, waiting-list controlled study. *Archives of General Psychiatry*, *60*(5), 517–522. http://dx.doi.org/10.1001/archpsyc.60.5.517.

Westra, H. A., Arkowitz, H., & Dozois, D. J. A. (2009). Adding a motivational interviewing pretreatment to cognitive behavioral therapy for generalized anxiety disorder: a preliminary randomized controlled trial. *Journal of Anxiety Disorders*, *23*, 1106–1117. http://dx.doi.org/10.1016/j.janxdis.2009.07.014.

Winchel, R. M., Jones, J. S., Molcho, A., Parsons, B., Stanley, B., & Stanley, M. (1992). The psychiatric institute trichotillomania scale (PITS). *Psychopharmacology Bulletin*, *28*, 463–476.

Woods, D. W., Flessner, C. A., Franklin, M. E., Keuthen, N. J., Goodwin, R. D., Stein, D. J., & Walther, M. R. (2006a). Trichotillomania Learning Center-Scientific Advisory Board. (2006). The Trichotillomania Impact Project (TIP): exploring phenomenology, functional impairment, and treatment utilization. *Journal of Clinical Psychiatry*, *67*, 1877–1888.

Woods, D. W., Wetterneck, C. T., & Flessner, C. A. (2006b). A controlled evaluation of acceptance and commitment therapy plus habit reversal for trichotillomania. *Behaviour Research and Therapy*, *44*(5), 639–656. http://dx.doi.org/10.1016/j.brat.2005.05.006.