Treatment of Trichotillomania With Behavioral Therapy or Fluoxetine

A Randomized, Waiting-List Controlled Study

Agnes van Minnen, PhD; Kees A. L. Hoogduin, MD, PhD; Ger P. J. Keijsers, PhD; Inge Hellenbrand, MA; Gert-Jan Hendriks, MD

Background: Both behavioral therapy (BT) and serotonin reuptake inhibitors have been reported effective in the treatment of trichotillomania. This study examines the efficacy of BT and fluoxetine hydrochloride compared with a waiting-list (WL) control group.

Methods: Forty-three patients with trichotillomania entered a 12-week randomized, WL-controlled study of BT and fluoxetine (60 mg/d). Forty patients (14 in the BT group, 11 in the fluoxetine group, and 15 in the WL group) completed the trial. Treatment effects were evaluated using the Massachusetts General Hospital Hairpulling Scale, and severity of hair loss was rated by independent assessors. In addition, we measured general symptoms of psychopathologic abnormalities and depression.

Results: For reducing the symptoms of trichotillomania, BT was superior. Patients in the BT group showed a significantly greater reduction in trichotillomania symptoms, higher effect sizes (Massachusetts General Hospital Hairpulling Scale: BT, 3.80; fluoxetine, 0.42; and WL, 1.09), and more clinically significant changes (BT, 64%; fluoxetine, 9%; and WL, 20%) than patients in the fluoxetine and WL groups. For severity of hair loss, a similar trend was also found in favor of the BT group. No significant differences between groups were established for general psychopathologic and depressive symptoms.

Conclusions: Behavioral therapy is highly effective for reducing symptoms of trichotillomania in the short term, whereas fluoxetine is not.

Arch Gen Psychiatry. 2003;60:517-522

Original Article

From the Department of Clinical Psychology, University of Nijmegen (Drs van Minnen, Hoogduin, and Keijsers and Ms Hellenbrand), and GGz Nijmegen (Dr Hendriks), Nijmegen, the Netherlands.
Powerful randomized controlled studies comparing the efficacy of BT and pharmacotherapy in the treatment of trichotillomania are still lacking. In this randomized, waiting-list (WL) controlled study in which we compared the treatment effects of BT and fluoxetine for trichotillomania, we tried to make up for this deficiency. We hypothesized that BT and fluoxetine are equally effective and are both more effective than the WL condition.

**METHODS**

**DESIGN**

Forty-three outpatients with trichotillomania were randomly assigned to 1 of 3 conditions: 6 sessions of BT during a 12-week period, a 12-week fluoxetine treatment (60 mg/d), or 12 weeks on a WL. The cell size (or number of patients per condition) needed for a statistical power of 80% for BT and WL was estimated at 8 based on a mean SD pretreatment-posttreatment improvement score on the severity scale of 8.6±4.2 and on the impairment scale of 3.3±2.3. This was determined according to the National Institute of Mental Health Trichotillomania Questionnaire used in a recent BT study with an estimated pretreatment-posttreatment correlation of 0.50 and a significance level of P<.05.

**PATIENTS**

Subjects were self-referred patients who had contacted a university outpatient clinic after seeing a national television show in which a BT program for trichotillomania was discussed.

Inclusion criteria were as follows: primary diagnosis of trichotillomania (DSM-IV criteria), age older than 16 years, and absence of organic brain disease, suicidal intent, or past or present psychosis. We excluded patients who were using antidepressants, patients who had previously used fluoxetine at a dose of 60 mg/d without result, patients who had previously been treated with BT comprising homework assignments, self-monitoring tasks, and interventions aimed at the stimulus-response chain, and women who were pregnant, trying to become pregnant, or lactating.

**PROCEDURE**

Eligible respondents received a standardized clinical interview in which the diagnostic criteria were verified (DSM-IV criteria), clinical features were established, and inclusion and exclusion criteria, including comorbidity, were checked. Subsequently, the study and rationale for BT or fluoxetine were explained, both orally and in writing. An hour later, a second standardized clinical interview was conducted by 2 of us to confirm the DSM-IV diagnosis and the inclusion and exclusion criteria. After receiving a detailed description of the study, patients gave their written informed consent and then participated in a pretreatment assessment. Next the assessor randomly assigned patients (by opening a sealed, numbered envelope) to 1 of the 3 treatment conditions. Randomization sequence was determined by referral sequence. After 12 weeks, a posttreatment assessment was conducted.

**TREATMENT**

**Behavioral Therapy**

Behavioral therapy consisted of 6 individual 45-minute, manual-based treatment sessions given once every other week. The treatment was aimed at self-control; using self-report and self-monitoring, the patients learned to control unwanted behavior in their own environment. The main elements were stimulus control (organizing the environment), stimulus-response interventions (interrupting the chain of response with other or incompatible activities), and response consequences (self-rewards). The therapist’s role consisted of behavior analysis, giving technical advice, and motivating the patient.

In session 1, the rationale of the treatment was explained and a behavior chain was constructed. For the duration of the treatment, the patient was given a daily homework assignment to write down every hour the number of hairs pulled and the total amount of time spent on hair pulling. The patient was also requested to save the pulled hairs in an envelope and give it to the therapist at the next session. In session 2, the results of the assignment were discussed and graphically displayed. Conscious awareness of hair-pulling behavior was increased by introducing several aids; for example, applying a strong perfume to the wrists or bandages around the fingers, or having the patient wear tinkling bracelets to signal that the arm was approaching the hair or that the patient had started pulling out hair. Additionally, most patients were instructed to put on gloves in high-risk situations (stimulus control). Besides increasing awareness, the wearing of gloves prevented patients from actual hair pulling. In session 3, stimulus-response interventions were selected such as going for a walk, calling a friend, or cleaning the kitchen. The patient was permitted to give in to the urge to pull the hair only after he or she had completed an activity, thereby postponing the unwanted behavior. Response consequences in the form of useful but tedious or unpleasant tasks were also introduced (eg, cleaning the bathroom or a 30-minute jog). Response consequences followed whenever hair pulling exceeded agreed levels. In sessions 4 and 5, the stimulus-response interventions and response consequences were discussed and extended. In the last session, relapse prevention was addressed.

**Fluoxetine**

Patients in the fluoxetine group received the drug at a dose of 60 mg/d for 12 weeks. Patients started with 20 mg/d, and during a 2-week period the dosage was titrated to 60 mg/d. To check and enhance compliance with treatment, 6 individual 30-minute sessions with the psychiatrist were provided comprising the following elements:

1. Patients were fully informed about the diagnosis, treatment rationale, and nature and course of possible adverse effects. We made sure that the patient adequately understood the information.

2. Patients were promised an alternative treatment in case of persistent (at least 4 weeks) or unexpected adverse events. Adverse drug reactions were checked using the adverse effect questionnaire developed by Fawcett et al, a semistructured instrument consisting of 7 categories that comprised psychological aspects, neurosensory and motor signs, neurovegetative signs, sexual disturbance, and 3 general categories. Before the start of the medication, none of the patients showed any of these symptoms. After 2 weeks of medication, 9 patients reported mild to moderate adverse effects: insomnia, drowsiness, fatigue, nausea, dry mouth, dizziness, excessive perspiration, tremor, headache, or delayed orgasm. At the end of the treatment, 7 patients still reported insomnia, fatigue, headache, excessive perspiration, weight loss, delayed orgasm, or anorgasmia. All adverse effects were described as mild except for the sexual effects, which were experienced as moderate to severe. In 1 patient, the fluoxetine dose was reduced during week 10 from 60 mg/d to 40 mg/d because of severe insomnia.
3. We stressed that patients could consult the prescribing psychiatrist by telephone during office hours if they had any worries or questions about the treatment or adverse effects.

4. Ample attention was paid to optimizing the therapeutic relationship by using motivational strategies to enhance the therapeutic bond and the patient’s motivation.

5. Appointment schedules were made in advance and were tailored to the patient’s wishes. At each visit, the next date and time were checked; in the event of a cancellation, the patient was contacted the same day to arrange a new appointment.

Note that no homework assignments were given so that contamination with elements of the BT condition could be prevented.

Waiting List

Patients in the WL condition did not receive any treatment for a period of 12 weeks. After this period, patients were randomly assigned to either BT or fluoxetine. The data for the latter treatments are not included in this study.

THERAPISTS

Therapists in the BT group were graduate students in clinical psychology who were fulfilling their practical training at the university outpatient clinic. They were adequately trained in the treatment program and were supervised weekly by 1 of us. Patients in the fluoxetine group were seen by an experienced psychiatrist.

ASSESSMENT

Assessors and Rating Periods

The pretreatment assessment was conducted before the randomization. The posttreatment assessment took place 2 weeks after the last treatment session, 12 weeks after the start of treatment. Assessors were blind to the patients’ treatment condition. In addition, for the BT and fluoxetine conditions, the Massachusetts General Hospital Hairpulling Scale (MGHHS) was completed before the start of each session.

Measures

The primary measure was the MGHHS.25 The MGHHS is a self-report measure assessing hair-pulling symptoms. It was developed using the Yale-Brown Obsessive-Compulsive Scale26,27 as a template. The scale consists of 7 items, each ranging from 0 (no symptoms) to 4 (extreme symptoms). In addition to the total score, the items related to actual hair pulling (items 4, 5, and 6) were added to constitute the actual-pulling subscale (MGHHS-AP). This scale has been demonstrated to have good psychometric properties.25,28 For our study the scale was translated into Dutch, in which the Cronbach α = .77 for the total scale and .68 for the actual-pulling subscale, reflecting adequate internal consistency.

Secondary measures included severity of hair loss, rated by the therapists. The assessors made pretreatment and posttreatment video recordings of the patient’s head and face using standard conditions and following standard instructions. For each patient, frontal, rear, and 2 side-view shots were made, each lasting 8 seconds. After receiving training and viewing video fragments twice each of 17 patients not included in this study, 2 independent raters assessed each videotape on a 4-point Likert scale that ranged from no loss of hair (1) to serious loss of hair (4). Videotapes were shown randomly with regard to pretreatment, posttreatment, and treatment condition. The raters were blind to the assessment and treatment condition, and each hair-pulling site was rated separately for every patient. Thus, when a patient pulled hair from both the scalp and the eyelashes, both sites were coded separately and independently. The Cohen κ was 0.76, indicating sufficient agreement between the 2 raters. In our study, the average assessment of both raters was used to establish severity of hair loss as rated by others. For patients with more than 1 hair-pulling site, we computed a composite score of all sites.

In addition, the Dutch version of the Symptom Checklist (SCL-90)29 was used, contains 90 items regarding general symptoms of psychopathologic abnormalities with a score range of 90 to 450, higher scores indicating more abnormalities. Internal consistency of the SCL-90 was shown to be good (Cronbach α = .89).

Finally, we used the Dutch version of the Beck Depression Inventory (BDI),30 a self-report measure of depressive symptoms consisting of 21 items with a score range of 0 to 63. Higher scores indicate more psychopathologic characteristics. Internal consistency of the BDI was good (Cronbach α = .85).

STATISTICS

For each of the 5 outcome measures, gain scores (pretreatment minus posttreatment scores) were calculated and entered into separate 1-way analysis of variance, which allowed us to distinguish the effects for the primary measures (MGHHS and MGHHS-AP) from those of the secondary measures (video ratings, SCL-90, and BDI). Analyses were conducted for the patients who completed the study and were repeated for the completers plus dropouts, for whom the last observation was carried forward. Because none of the patients for whom video ratings were available were noncompleters, no such analyses were conducted with these video ratings.

Effect sizes were calculated using the effect size index (Cohen d) for repeated measures.31 The measure of clinically significant change was used in addition to statistical significance tests and effect sizes because the latter measures are limited; they provide no information on the variability of response to treatment within the sample and provide little information on the efficacy of treatment. Clinically significant change was determined following the suggestions of Jacobson and Truax32 for calculating the cutoff point, XT, as follows:

\[ X_T = \frac{\text{pretreatment} - 2S_1 - S_e}{2}, \]

where \( S_e \) is the SD of pretreatment MGHHS scores for all patients. When a patient’s posttreatment score is more than 2 SDs plus half of the MGHHS error score lower than the pretreatment score for all patients, Jacobson and Truax consider the patient to have achieved a clinically significant change. The cutoff score in our sample was as follows:

\[ X_T = 15.86 - 2 \times 4.39 - 0.75/2 = 6.70. \]

RESULTS

PATIENT FLOW

Fifty patients met the inclusion criteria and were offered treatment. Seven of these patients refused the randomization procedure and were excluded from the study. Of the 43 patients who started treatment within the study, 15 were randomly assigned to the BT group, 13 to the fluoxetine group, and 15 to the WL group. One patient dropped out of BT because of the long commute to the clinic. Two patients dropped out of fluoxetine treatment before week 12: one patient because of the long com-
Table

<table>
<thead>
<tr>
<th>Pretreatment and Posttreatment Scores, Pretreatment-Posttreatment Correlations, and Effect Sizes for the Behavioral Therapy, Fluoxetine, and Waiting-List Conditions for the MGHHS (Total Scale and Actual-Pulling Subscale), Hair Loss (Video Ratings), BDI, and SCL-90*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Behavioral therapy</td>
</tr>
<tr>
<td>MGHHS, total</td>
</tr>
<tr>
<td>MGHHS-AP</td>
</tr>
<tr>
<td>Hair loss (video ratings)</td>
</tr>
<tr>
<td>BDI</td>
</tr>
<tr>
<td>SCL-90</td>
</tr>
<tr>
<td>Fluoxetine</td>
</tr>
<tr>
<td>MGHHS, total</td>
</tr>
<tr>
<td>MGHHS-AP</td>
</tr>
<tr>
<td>Hair loss (video ratings)</td>
</tr>
<tr>
<td>BDI</td>
</tr>
<tr>
<td>SCL-90</td>
</tr>
<tr>
<td>Waiting list</td>
</tr>
<tr>
<td>MGHHS, total</td>
</tr>
<tr>
<td>MGHHS-AP</td>
</tr>
<tr>
<td>Hair loss (video ratings)</td>
</tr>
<tr>
<td>BDI</td>
</tr>
<tr>
<td>SCL-90</td>
</tr>
</tbody>
</table>

Abbreviations: BDI, Beck Depression Inventory; MGHHS, Massachusetts General Hospital Hairpulling Scale; MGHHS-AP, actual-pulling subscale of the MGHHS; SCL-90, Dutch version of the Symptom Checklist.

*Data are presented as mean ± SD unless otherwise indicated.

mute, and the other patient as a result of serious injuries sustained in a road accident. Patients were recruited between January 1999 and July 1999. Treatments were conducted between August 1999 and April 2000.

Five patients (11.6%) were men, and 38 were women (BT, 3 men; fluoxetine, 1 man; WL, 1 man). The mean ± SD age was 31.9 ± 11.5 years, ranging from 17 to 57 years (BT, 31.7 ± 11.9 years; fluoxetine, 33.7 ± 8.2 years; WL, 30.5 ± 13.7 years). The mean ± SD duration of symptoms prior to treatment was 19.1 ± 12.2 years (BT, 20.8 ± 15.3 years; fluoxetine, 19.0 ± 8.4 years; WL, 17.6 ± 12.0 years). Seventeen patients (39.5%) pulled hair from the scalp (BT, 5; fluoxetine, 3; WL, 9), and 9 patients (20.9%) pulled hair from their eyebrows, eyelashes, or both (BT, 2; fluoxetine, 4; WL, 3). The remaining 17 patients pulled hair from the scalp, eyelashes, and/or eyelashes in various combinations. Eleven of them pulled hair from other parts of the body as well. In addition to trichotillomania, some patients had mood disorders (BT, 3; fluoxetine, 3; WL, 2), anxiety disorders (BT, 1; fluoxetine, 0; WL, 0), or other impulse control disorders (BT, 1; fluoxetine, 2; WL, 0). With regard to comorbidity, no significant differences between groups were observed. The Table lists the mean pretreatment and posttreatment scores of the primary and secondary outcome measures for the BT, fluoxetine, and WL groups separately.

At pretreatment assessment, there were no significant differences between the 3 groups with respect to the MGHHS (F$_{2,37}$=0.37; P=.69), MGHHS-AP (F$_{2,40}$=0.83; P=.44), BDI (F$_{2,40}$=0.76; P=.47), or SCL-90 (F$_{2,38}$=0.36; P=.70).

For the MGHHS gain score, there was a main effect for group (F$_{2,37}$=12.32; P<.001) indicating that the improvement was different among groups. Post-hoc tests with the Bonferroni correction showed that the MGHHS gain scores in the BT group were significantly larger than those in the fluoxetine group (P<.001) and the WL group (P<.001).

Similar results were found for the MGHHS-AP. There was a main effect for group (F$_{2,37}$=17.39; P<.001), signifying that improvement was different for each group. Post-hoc tests with the Bonferroni correction showed that the MGHHS-AP gain scores in the BT group were significantly larger than those in the fluoxetine group (P<.001) and the WL group (P=.004). In addition, MGHHS-AP gain scores in the WL group proved to be significantly larger than in the fluoxetine group (P=.04).

The video ratings revealed a trend toward significance for effects between groups (F$_{2,39}$=3.23; P=.06). Post-hoc analyses with the Bonferroni correction showed that the gain scores in the BT group tended to be significantly larger than in the WL group (P=.09). No significant main effects for group were found for the BDI (F$_{2,37}$=0.80; P=.46) or SCL-90 (F$_{2,35}$=0.28; P=.76) gain scores.

The previously mentioned results were obtained for the patients who completed the study. Comparable results were obtained after repeating the analysis for all patients, including dropouts, for whom the last observations were carried forward.

To gain further insight into the clinical significance of the improvements achieved in each of the 3 groups, effect sizes were calculated for all outcome measures (Table). On the MGHHS total scale, effect sizes were found to be high for the BT group (3.80) and moderate to low for the fluoxetine (0.42) and WL (1.09) groups. Regarding the other outcome measures, effect sizes in the BT group proved superior to the other groups.

For the primary outcome measure, the MGHHS, the degree of clinically significant change was established. Patients with an MGHHS posttreatment score...
lower than 6.70 achieved a clinically significant change. This pertained to 9 (64%) of 14 patients in the BT group, 1 (9%) of 11 patients in the fluoxetine group, and 3 (20%) of 15 patients in the WL group. These differences proved to be significant ($\chi^2_{1.10} = 10.26; P = .006$).

**COMMENT**

The aim of this controlled study was to investigate the treatment effects of BT and fluoxetine after 12 weeks of treatment. The results are consistent in showing that BT is superior to both fluoxetine treatment and postponement of treatment (WL condition). The effect size for the MGHHS was high (3.80), and 64% of the patients achieved a clinically significant change. Patients in the BT group showed significant reductions in the severity of trichotillomania symptoms, including actual hair pulling. When the hair loss was rated by others using videotapes, patients who had undergone BT were found to have improved significantly. No significant differences were found between BT, fluoxetine treatment, and the WL condition with respect to reduction of depressive symptoms (BDI) or general symptoms of psychopathologic abnormalities (SCL-90). For severity of hair loss as rated by others, a trend toward a significant difference between groups was found, supporting the effectiveness of BT. Our favorable treatment results for BT are in line with the results of previous controlled and open-trial studies.1-5,33

Fluoxetine treatment was shown to be ineffective in reducing trichotillomania symptoms. The low improvement rates for fluoxetine treatment found in this study confirm the results of 2 previous controlled studies15,16 but contradict those from another controlled study14 and from several open-trial studies.9-12 One reason for the contradictory effects of fluoxetine treatment in trichotillomania may be that dosages and treatment durations differ in the various studies. It may be argued that the dose (60 mg/d) in our study was too low or that the treatment duration of 12 weeks was too short. However, there are no clear indications in the literature to corroborate this. Christenson et al15 similarly found no treatment effects for fluoxetine in 16 patients with trichotillomania using dosages that varied between 20 mg/d and 80 mg/d during a 12-week period. Streichenwein and Thornby16 also failed to demonstrate a significant improvement even after prolonging their treatment for another 12 weeks and maintaining a fixed dose of 80 mg/d for 6 weeks. They concluded that serotonin may not be the main dysfunctional neurotransmitter in trichotillomania.

The fact that fluoxetine treatment was unsuccessful in our study may also be because, contrary to common practice, we kept no daily record of hair pulling for that condition to differentiate fluoxetine treatment from BT. Daily records may enhance patients’ awareness of hair pulling. In the study by Winchel et al,12 for example, patients generally attributed their improvements after fluoxetine treatment to becoming more consciously aware of their hair pulling.

Remarkably, in this study the patients in the WL condition for 12 weeks showed a significant reduction in hair-pulling symptoms. The effect sizes for hair pulling (MGHHS, 1.09; MGHHS-AP, 1.39) were good. The improvement in the WL condition may be attributed to expectancy effects because these patients had been promised treatment after 12 weeks. Alternatively, during the intake and assessment procedures, all patients received information about trichotillomania and its treatment possibilities; some patients may have picked up treatment suggestions from the information leaflets. Furthermore, they also met other patients with trichotillomania in the waiting room.

Similar possible expectancy effects or influence of treatment suggestions seem to be absent in the fluoxetine condition. Perhaps, compared with the WL patients, those in the fluoxetine group felt less need to consider the treatment suggestions in the information leaflets. In sum, we found that BT was superior to fluoxetine in the treatment of trichotillomania. However, several remarks concerning this study need to be made. First, our primary measure to assess the severity of trichotillomania symptoms was based on self-report. Although to date the MGHHS is the only self-report measure for which good psychometric properties have been established, self-reports have limitations. We therefore opted for a supplementary, more objective hair loss measure; namely, independent ratings of videotapes showing the patients’ affected sites. However, this technique presented various problems. Several patients were so embarrassed about their hair loss that they refused to be videotaped, resulting in missing values. Also, some hair-pulling sites (eg, eyelashes) were difficult to discern on the videotape, and in some patients, hair pulling may have occurred in a distributed fashion. In addition, the degree of hair regrowth may have varied among patients and hair-pulling sites, implying that a reduction in hair pulling may not be directly related to an observed increase in hair growth. Because of these assessment difficulties, the results regarding the video ratings should be interpreted with caution. Second, our results are based on a relatively small number of self-referred patients who contacted the clinic in response to a television show mentioning BT, which may restrict the generalizability of the study. However, the estimated statistical power was satisfactory for comparing the main treatment effects for the primary measures, and the baseline characteristics of our patients were highly comparable with those of patients who had been officially referred to a trichotillomania hospital.35 Finally and most important, only the immediate posttreatment effects have been addressed in this study. Whether the superior treatment effects for BT will be maintained in the long term is unknown. Follow-up data from other studies of BT have shown that relapse rates are high,3,35 illustrating the importance of this issue. Follow-up studies are essential to provide information about the stability of symptom reduction. Presently, we are collecting follow-up data to address several of these issues and to learn more about variables that may influence the effectiveness of BT.36

To conclude, in our study BT was shown to be highly effective in reducing symptoms of trichotillomania at posttreatment assessment. The challenge for the future is to further refine this treatment to obtain long-term improvement.

Submitted for publication December 14, 2001; final revi-
REFERENCES