Pain 121 (2006) 181–194

PAIN

www.elsevier.com/locate/pain

Research papers

Short- and long-term efficacy of brief cognitive-behavioral therapy for patients with chronic temporomandibular disorder pain: A randomized, controlled trial

Judith A. Turner [a,b,*], Lloyd Mancl [c], Leslie A. Aaron [d]

[a] Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, WA, USA
[b] Department of Rehabilitation Medicine, University of Washington School of Medicine, Seattle, WA, USA
[c] Department of Dental Public Health Sciences, University of Washington School of Dentistry, Seattle, WA, USA
[d] Department of Oral Medicine, University of Washington School of Dentistry, Seattle, WA, USA

Received 26 August 2005; received in revised form 9 November 2005; accepted 21 November 2005

Abstract

We evaluated the short- and long-term efficacy of a brief cognitive-behavioral therapy (CBT) for chronic temporomandibular disorder (TMD) pain in a randomized controlled trial. TMD clinic patients were assigned randomly to four sessions of either CBT ($n = 79$) or an education/attention control condition ($n = 79$). Participants completed outcome (pain, activity interference, jaw function, and depression) and process (pain beliefs, catastrophizing, and coping) measures before randomization, and 3 (post-treatment), 6, and 12 months later. As compared with the control group, the CBT group showed significantly greater improvement across the follow-ups on each outcome, belief, and catastrophizing measure (intent-to-treat analyses). The CBT group also showed a greater increase in use of relaxation techniques to cope with pain, but not in use of other coping strategies assessed. On the primary outcome measure, activity interference, the proportion of patients who reported no interference at 12 months was nearly three times higher in the CBT group (35%) than in the control group (13%) ($P = 0.004$). In addition, more CBT than control group patients had clinically meaningful improvement in pain intensity (50% versus 29% showed $\geqslant$50% decrease, $P = 0.01$), masticatory jaw function ($P < 0.001$), and depression ($P = 0.016$) at 12 months (intent-to-treat analyses). The two groups improved equivalently on a measure of TMD knowledge. A brief CBT intervention improves one-year clinical outcomes of TMD clinic patients and these effects appear to result from specific ingredients of the CBT.

Keywords: Cognitive-behavioral therapy; Temporomandibular disorders; Chronic pain; Brief psychotherapy; Randomized controlled trial

1. Introduction

Cognitive-behavioral therapies (CBT) aim to decrease maladaptive, and to increase adaptive, patient cognitions and behaviors. These treatments are effective for a variety of chronic pain problems (Keefe and Caldwell, 1997; Morley et al., 1999; Astin et al., 2002; Eccleston et al., 2002; Weydert et al., 2003; Chen et al., 2004). CB therapies are time-limited; the median number of treatment hours was 16 in the randomized controlled trials (RCTs) of CBT for chronic pain included in a comprehensive review (Morley et al., 1999). However, limitations in insurance coverage for psychological therapies and the fact that most psychotherapy clients attend fewer than six sessions (Shapiro et al., 2003) create the need for effective very brief (six or fewer sessions) CBTs for chronic pain problems.

We previously reported changes on electronic daily diary measures of outcome and therapy process variables over the course of treatment in an RCT of a very

* Corresponding author. Tel.: +1 206 543 3997; fax: +1 206 685 1139.
E-mail address: jturner@u.washington.edu (J.A. Turner).

0304-3959/$32.00
doi:10.1016/j.pain.2005.11.017

brief CBT for patients with chronic temporomandibular disorder (TMD) pain (Turner et al., 2005). TMDs are a group of conditions that involve the temporomandibular joints (TMJ), masticatory muscles, and associated structures, and share the common symptoms of pain, restricted jaw function, and TMJ noises (Dimitroulis, 1998). The etiology of the most common types of TMD is largely unknown (Dimitroulis, 1998). TMDs are the most frequent facial pain problems (Dworkin, 1995b), with an estimated prevalence of 10–12% (Von Korff et al., 1988; Dworkin et al., 1990a; LeResche, 1997). Patients with TMD pain are similar to patients with other chronic pain conditions (e.g., headache and back pain) in terms of pain intensity and associated behavioral and psychological dysfunction (Dworkin, 1995a), pain persistence and recurrence (Dworkin et al., 1989), and refractoriness of pain to treatment (Rudy and Turk, 1995). Suggesting the importance of psychosocial factors in TMD problems, individuals with TMD vary widely in levels of disability (e.g., pain interference with customary activities; difficulty with masticatory and non-masticatory jaw activities such as opening jaws to bite food, chewing, and kissing) and psychosocial dysfunction (Butterworth and Deardorff, 1987; Rudy et al., 1989; Suvinen et al., 1997) but objective findings do not appear to underlie these differences (Rudy et al., 1989; Dworkin, 1995a), and changes over time in jaw function measures are not clearly related to course of pain (Ohrbach and Dworkin, 1998).

We now report, following the Consolidated Standards of Reporting Trials (CONSORT) guidelines (Moher et al., 2001), the short- and long-term results of our RCT of CBT versus an education/attention control condition for chronic TMD pain as assessed by standard questionnaire measures of outcome and therapy process variables. We hypothesized that the CBT group would show greater short- and long-term improvement on each outcome (pain, activity interference, jaw functioning, and depression) and process (pain-related beliefs, catastrophizing, and coping) measure. We examined clinically important as well as statistically significant change on the outcome measures. The primary outcome was activity interference and the primary endpoint was one-year.

2. Methods

2.1. Setting and participants

The study was approved by the University of Washington (UW) Institutional Review Board and all participants provided written informed consent. Study participants were recruited by research staff from patients seeking care at the UW Orofacial Pain Clinic between June 2001 and February 2004. Study inclusion criteria were: (1) age 18 years or older; (2) a Research Diagnostic Criteria/Temporomandibular Disorders (RDC/TMD) Axis I TMD diagnosis (Dworkin and LeResche, 1992) made by an oral medicine specialist based on a structured RDC/TMD clinical examination; (3) residence within a 2-h drive of the TMD clinic; (4) facial pain for at least three months; (5) facial pain-related disability, as defined by a chronic pain grade (Von Korff et al., 1992) of II high, III, or IV (see Section 2.3 for definitions); and (6) ability to communicate in English. Study exclusion criteria (assessed by the patient's oral medicine specialist and the study coordinator) were need for further diagnostic evaluation, pending litigation or disability compensation for pain, current or previous CBT for pain, and major medical or psychiatric conditions that would interfere with ability to participate.

Among the 366 patients approached and found to be eligible for the RCT, 158 (43%) enrolled (Fig. 1). The most commonly cited reason for decision not to enroll was the time commitment required to attend four sessions (most patients worked and many patients in this referral clinic lived a 1- to 2-h drive from the clinic). Two participants were withdrawn from the study shortly after randomization because severe psychiatric problems undetected before enrollment (and that would have made them ineligible for the study had they been known) became evident. The other 156 subjects randomized did not differ significantly from patients who declined to participate in the study ($n = 208$) in gender, race, education, chronic pain grade, characteristic pain intensity, activity interference, or any RDC/TMD clinical diagnosis (see Measures for definitions). However, study participants were somewhat younger on average [mean (SD) = 37.0 (11.4) versus 39.6 (12.5) years; $P = 0.04$]. All participants who completed at least one post-treatment or follow-up assessment were included in analyses for this report. Those randomized who did ($n = 148$, 95%) versus did not ($n = 8$, 5%) complete at least one post-treatment or follow-up assessment did not differ significantly at baseline on any sociodemographic variable assessed, chronic pain grade, RDC/TMD diagnosis, pain duration, or any outcome or process measure, with one exception: those who provided no follow-up data had higher baseline Chronic Pain Coping Inventory (CPCI) Coping Self-Statements scale (see Section 2.3) scores (that is, greater reported use of coping self-statements) [mean (SD) = 3.7 (2.1) versus 2.5 (1.7), $P = 0.049$].

2.2. Procedures

Participants were asked to complete questionnaires at home and return them in person or by mail prior to randomization (baseline), and 3 (post-treatment), 6, and 12 months after randomization. Participants were compensated for questionnaire completion (post-treatment $25, 6-month $10, 12-month $50, and an additional $15 if all questionnaires were completed on schedule). Participants were also asked to complete electronic diaries three times daily for eight weeks while participating in their randomly assigned treatment.

All study participants received treatment as usual from their dentist at the Orofacial Pain Clinic. These treatments were conservative and typically included instruction in jaw posture monitoring and correction (including instruction to keep jaws relaxed and teeth apart, but no training in muscle relaxation techniques), advice to apply heat and/or cold to painful facial areas, and recommendations concerning diet modifications. Medications (e.g., non-steroidal anti-inflammatory drugs), jaw stretching exercises, and occlusal splints were prescribed for some patients.

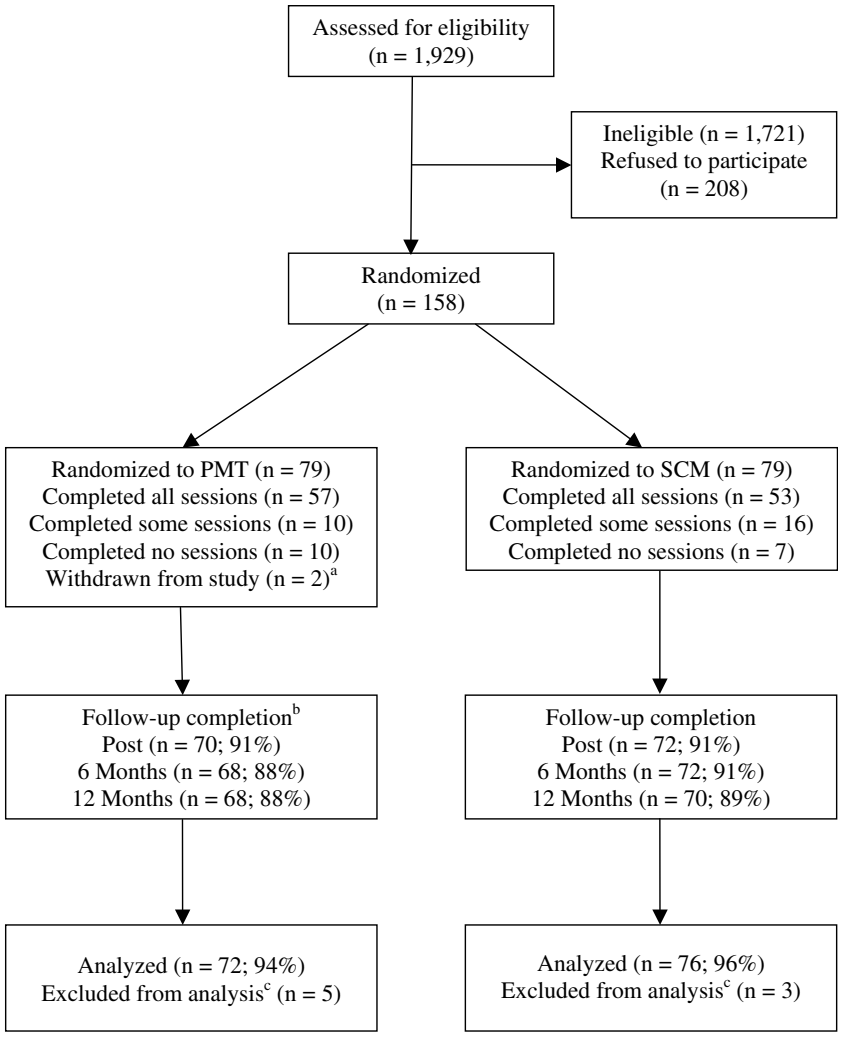

Fig. 1. Participant flow through the RCT. [a]Withdrawn for psychiatric reasons. [b]Follow-up completion rates based on $n = 77$ randomized (not withdrawn). [c]Excluded from analysis due to no follow-up data, although included in the multiple imputation analyses.

2.3. Measures

2.3.1. Outcome measures

2.3.1.1. Activity interference and pain intensity. The Graded Chronic Pain Scale (GCPS) (Von Korff et al., 1992; Von Korff, 2001) was used to assess pain intensity and interference with usual daily activities. The primary outcome, *activity interference* (Von Korff, 2001), was calculated by averaging 0–10 ratings of pain interference with daily activities, work/housework activities, and recreational/social activities in the past month. *Characteristic pain intensity* was calculated by averaging 0–10 ratings of current pain and average and worst pain in the past month (Dworkin et al., 1990b; Von Korff et al., 1992; Von Korff, 2001). The characteristic pain intensity and activity interference scores have good internal consistency, test–retest reliability, and validity (Underwood et al., 1999; Von Korff, 2001). The GCPS also allows individuals to be classified into five chronic pain grades: 0 = no pain, I = low pain intensity and low pain-related disability, II low = high pain intensity and no pain-related disability, II high = high pain intensity and low pain-related disability, III = moderate pain-related disability, and IV = severe pain-related disability.

2.3.1.2. Jaw use limitations. The Mandibular Function Impairment Questionnaire (MFIQ) (Stegenga et al., 1993b) is a 17-item measure with two subscales (masticatory and non-masticatory jaw disability) demonstrated to be sensitive to change with treatment for TMD (Stegenga et al., 1993a). Scores on each subscale have a possible range of 0–1. Categories of low, moderate, and severe mandibular function impairment have been developed for the MFIQ (Stegenga et al., 1993b).

2.3.1.3. Depression. The 21-item Beck Depression Inventory (BDI) (Beck and Beamesderfer, 1974; Beck et al., 1979) was used to assess depressive symptom severity. The BDI has high internal consistency, adequate test–retest reliability, and validity (Beck et al., 1988), and is a valid screening instrument for depression among patients with chronic pain (Turner and Romano, 1984; Love, 1987; Geisser et al., 1997). Some investigators have found more than one underlying factor in the

BDI responses of patients with chronic pain and have suggested that it may be important to analyze item endorsement patterns in this population (Morley et al., 2002). However, we used standard scoring procedures for purposes of examining overall change in depressive symptom severity during the study, because removal of the somatic items does not appear to improve the BDI's accuracy (Geisser et al., 1997). There is general agreement that scores of 21 or higher suggest at least a moderate level of depressive symptoms (Kendall et al., 1987; Geisser et al., 1997; Rogers et al., 2005); therefore, we chose this cutpoint as an indicator of clinically meaningful depression in the study participants.

2.3.2. Process measures

2.3.2.1. Pain beliefs. We administered three scales from the Survey of Pain Attitudes (SOPA) (Jensen et al., 1994): Disability (belief that one's pain is disabling), Harm (belief that pain signifies damage and that activity should be avoided), and Control (belief in one's personal control over pain). These scales have good test–retest stability, validity, and internal consistency (Jensen and Karoly, 1992; Strong et al., 1992; Jensen et al., 1994). Scores on each scale can range from 0 to 4, with higher scores indicating greater agreement with the belief.

Participants also completed the 8-item TMD Self-Efficacy Scale (SES), which is a modification (by replacing the word 'arthritis' with 'facial pain') of the Arthritis Self-Efficacy Scale (Gonzalez et al., 1995; Lorig et al., 1996). On this measure, patients rate on a scale numbered from 0 = 'very uncertain' to 10 = 'very certain' their certainty that they can decrease their pain quite a bit, keep facial pain from interfering with their sleep, keep their pain from interfering with the things they want to do, regulate their activity so as to be active without aggravating their pain, keep the fatigue caused by pain from interfering with the things they want to do, do something to feel better if they are feeling blue, manage facial pain during their daily activities, and deal with the frustration of facial pain. Scale scores are calculated as the mean of the eight ratings, with higher scores indicating greater self-efficacy. We previously reported that this scale had excellent internal consistency (Cronbach's $\alpha = 0.91$) and validity in the sample of TMD patients enrolled in this study (Brister et al., in press). Although the SOPA Control scale and the TMD SES are moderately correlated ($r = 0.54$ in this sample), the former scale assesses solely the belief in ability to control one's pain, whereas the SES assesses confidence in ability not only to decrease pain but also to manage specific pain-related problems.

Pain catastrophizing was assessed by two scales. The CSQ Catastrophizing scale has excellent internal consistency (Rosenstiel and Keefe, 1983; Keefe et al., 1989) and has been shown to be associated with various measures of functioning in samples of patients with different pain conditions (Keefe et al., 1987, 1989; Jensen and Karoly, 1991; Dozois et al., 1996; Martin et al., 1996), including TMD (Turner et al., 2001). Scores can range from 0 to 6, with higher scores indicating greater catastrophizing. The four-item Rumination subscale of the Pain Catastrophizing Scale (PCS) captures aspects of catastrophizing not assessed by the CSQ: ruminative thoughts, worry, and an inability to inhibit pain-related thoughts (Sullivan et al., 1995). The Rumination scale is associated significantly with pain and disability (Sullivan et al., 1998; Osman et al., 2000), and has good internal consistency and discriminant validity (Osman et al., 2000). Scores on the scale can range from 0 to 16, with higher scores indicating greater tendency to ruminate about pain.

2.3.2.2. Pain coping. We administered four scales from the Chronic Pain Coping Inventory (CPCI), a measure of strategies used in the past week to cope with chronic pain (Jensen et al., 1995). One of the scales, Rest, measures a pain coping response that is believed to be related to worse adjustment of patients with chronic pain and that is typically targeted for reduction in cognitive-behavioral treatments (Jensen et al., 1995). The other three scales, Task Persistence, Coping Self-Statements, and Relaxation, are believed to be adaptive coping responses and are typically encouraged in cognitive-behavioral treatments (Jensen et al., 1995). The CPCI scales have demonstrated internal consistency, test–retest reliability, and validity (Jensen et al., 1995). Scores on each scale can range from 0 to 7, with higher scores indicating greater use of the coping strategy.

2.3.3. Measures to assess treatment credibility, TMD knowledge, and treatment helpfulness

As a check on the two study interventions' equivalence in factors other than active ingredients of CBT, we administered a measure of treatment credibility at baseline and a measure to assess knowledge about TMD at baseline and post-treatment. We also administered a measure at post-treatment to assess study participants' views of the helpfulness of treatment components.

2.3.3.1. Treatment credibility. Treatment credibility questions (Borkovec and Nau, 1972) adapted for use in studies of psychological treatments for various pain and other problems (Bradley et al., 1987; Keefe et al., 1990; Safren et al., 1997; Addis et al., 2004) were also adapted for use in this study. At baseline, prior to randomization, study participants were given a brief description of the two study interventions. Participants rated on 0–10 scales (0 = not at all, 10 = extremely) how logical each treatment seemed to them, how confident they were that each treatment would help them better control their pain, how confident they were that each treatment would decrease their pain, and how confident they would be in recommending each treatment to a friend with similar problems. These ratings were averaged to create a single *treatment credibility* score for each treatment. The treatment credibility measures had high internal consistency in this sample (Cronbach's $\alpha = 0.92$ for credibility of the control condition and 0.93 for credibility of the CBT condition).

2.3.3.2. TMD knowledge. Participants were asked whether they agreed or disagreed with 10 statements about TMD (e.g., 'All individuals with clicking jaw joints should have treatment,' 'TMD affects many more women than men') used in previous research with TMD patients (Dworkin et al., 1994). Articles included in both the CBT and control condition participant manuals (described in Section 2.5) provided information bearing on each question. A score reflecting the proportion of questions answered correctly was calculated.

2.3.3.3. Treatment helpfulness. At the post-treatment assessment, study participants rated on 0–10 scales (0 = not at all helpful, 10 = extremely helpful) the helpfulness of four components shared by the two treatments: information about

TMD and TMD treatments, instructions for correcting jaw posture and habits, jaw stretching exercises (for patients whose dentists recommended their use), and reading material. Participants also used 0–10 scales (0 = not at all, 10 = extremely) to rate overall treatment program helpfulness and their satisfaction with the program.

2.4. Randomization

Study participants were assigned randomly to four individual biweekly sessions over eight weeks of either CB pain management training (PMT) or an education/attention control condition (self-care management; SCM). Randomization was stratified by participant chronic pain grade and gender to ensure that the two groups would be nearly balanced on these characteristics. Randomization assignments were generated by a biostatistician (LM) using randomly selected block sizes of two or four using the sample function of the S-PLUS® statistical software (Insightful Corporation, Seattle, WA) to prevent determination of the treatment assignment. Treatment assignments were recorded on slips of paper numbered consecutively within each stratum and sealed in envelopes sequentially numbered by stratum. Randomization assignment was concealed to all study personnel until envelopes were opened by research staff after subject consent was obtained.

2.5. Intervention protocol

Participants in each study condition were given a manual with materials to read between sessions and discuss in sessions. Both the PMT and the SCM manuals included an article about TMDs and TMD treatments and instructions for TMD self-care activities (e.g., monitoring jaw posture, relaxing jaws, avoiding certain jaw movements and activities, application of heat or cold, diet modifications, and jaw stretching exercises to do if prescribed by the patient's dentist). In both study conditions, the treatment provider made brief (<15 min) telephone calls to patients in weeks between sessions to inquire about homework assignment completion and address questions, and at 2, 4, 8, 12, 16, 20, and 24 weeks after the fourth session to ask how the patient was doing and whether he/she had any questions. In the PMT condition, the treatment provider also inquired in the telephone calls about the patient's use of TMD self-care and coping strategies emphasized in the PMT sessions and progress towards goals identified in the sessions. In both conditions, the treatment provider followed a detailed written protocol in each session and telephone call, and checked off each protocol activity as it was delivered. Although all study participants were encouraged to attend all four sessions, some participants failed to attend one or more sessions (numbers summarized in Section 3). In such cases, material in missed sessions was covered to the extent possible in subsequent telephone calls and sessions.

2.5.1. PMT

PMT participants were seen by one of three licensed clinical psychologists, each of whom had prior experience conducting CBT, including the techniques in the PMT protocol, with patients with chronic pain. Two of the psychologists were trained and supervised (in regular meetings throughout the study) in the PMT protocol by the third psychologist (JAT), who has had over 25 years' experience conducting CBT with patients with chronic pain, including patients with TMD. The treatment was based on standard CB pain therapies (Turner and Romano, 2001) and a previously studied CB intervention for chronic TMD pain (Dworkin et al., 2002). The PMT patient manual included articles concerning psychological aspects of pain, challenging negative thoughts about pain, relaxation, and other behavioral techniques for pain management, coping with pain flare-ups, and relapse prevention.

At each session, patients completed a 'personal TMD health care plan' for activities to complete between sessions. Patients checked off activities as they completed them on a daily basis, then brought the completed plan to each session for discussion. Certain activities were recommended to all PMT participants (e.g., check and correct jaw posture, progressive relaxation practice, and breathing exercises) and other activities were added as part of working toward specific patient-identified activity goals (e.g., increasing physical exercise such as walking). Jaw stretching exercises were included if they had been prescribed by the patient's dentist. The psychologist helped each patient identify potential obstacles to completing activities and possible solutions.

Each PMT session included instruction and practice in progressive relaxation and abdominal/diaphragmatic breathing techniques (Bernstein et al., 2000; Syrjala, 2001). Participants were given a relaxation audiotape and asked to practice relaxation daily. The PMT treatment also included discussion of fear-avoidance (Waddell et al., 1993; Vlaeyen and Linton, 2000), training in the identification and challenging of negative thoughts in response to pain (Turner and Romano, 2001), and discussion of relapse prevention and ways to maintain gains and deal with setbacks (Turner and Romano, 2001).

2.5.2. SCM

The SCM condition was designed to control for the effects of natural history/time, TMD education, patient expectations, completing study measures, and attention. The SCM condition did not include specific CBT techniques. The intervention was conducted by bachelor's level TMD patient educators trained and supervised (in weekly sessions, including review of the session protocol checklists) by a licensed clinical psychologist. The educator focused only on the SCM structured protocol content in each session and did not give any advice or recommendations to study participants beyond that in the protocol. In addition to the information about TMD described above, the SCM patient manual included general health care information (e.g., about pain medications, communicating with health care providers, and making treatment decisions). The sessions focused on reviewing the main points of each article in the manual and discussing the patient's reactions and questions. If patients raised specific questions about their TMD problems or treatments, the educator advised the patient to discuss these with his/her dentist.

2.6. Statistical power

A priori power calculations for this trial were based on data from previous studies of UW Orofacial Pain Clinic patients. We expected the decrease in activity interference from baseline to the 12-month follow-up to be 35–40% in the SCM group and 60–65% in the PMT group. Based on estimates of mean

(SD) baseline activity interference derived from our previous studies, a sample size of 70 per group could detect a difference (estimated to be 1.4 points) between groups in mean activity interference (the primary outcome) at the 12-month follow-up, assuming a standard deviation of 2.6, with 88% power for a two-sided *t*-test with a 0.05 significance level. This sample size also allowed sufficient statistical power (92%) to detect a difference (estimated to be 1.3 points) between groups in characteristic pain intensity at 12-month follow-up with an assumed standard deviation of 2.2.

2.7. Statistical analysis

Intent-to-treat analyses included all randomized participants for whom follow-up data were available (72 PMT and 76 SCM participants; 94% and 96%, respectively, of those randomized to PMT and SCM). To assess comparability of the two study groups at baseline on sociodemographic, TMD, and questionnaire measure variables, we used *t*-tests (for continuous and ordinal variables) and χ^2 tests (for categorical variables). To compare the two study groups on the outcome and process measures over time, we fit linear regression models, controlling for baseline values of the dependent variable examined and using generalized estimating equations (GEE) to adjust for possible correlation within patients over the assessment time points (Diggle et al., 2002). The two study groups did not differ at baseline on any sociodemographic or outcome variable (see Section 3). However, the groups did differ on two clinical TMD diagnoses (see Section 3), and although neither diagnosis was significantly related to any of the outcome measures, we performed sensitivity analyses adjusting for the two TMD diagnoses.

We also performed two sensitivity analyses to assess whether results were affected by participants' pretreatment views of the credibility of the study treatments. First, we repeated the regression analyses, this time controlling for participants' ratings of the credibility of the treatment to which they were subsequently randomized (i.e., we controlled for PMT participants' ratings of the credibility of PMT and for SCM participants' ratings of the credibility of SCM). A finding of significant differences between treatments even after controlling for the credibility of the assigned treatment would suggest that differences between treatments are unlikely to be due solely to patients' pretreatment expectations. However, such an analysis may result in the statistical problem of the confounding of treatment credibility ratings and treatment assignment, which could severely attenuate a true treatment effect, if present. Therefore, a more appropriate statistical analysis might be to enter two scores for each participant: their rating of the credibility of SCM and their rating of the credibility of PMT. Thus, we conducted the regression analyses again, controlling for both treatment credibility scores.

Examination of the distributions of the outcome and process measures revealed considerable skewness for activity interference and catastrophizing at the follow-up assessments. Substantial proportions of study participants had scores of 0 on these measures, and no transformation of scores could make the distribution shapes normal and similar for the two groups. We therefore dichotomized scores on these measures into categories of 0 versus some and performed logistic regression analyses using GEE to compare the two study groups on these measures over time. The distribution of BDI scores was also skewed, but normalized by a square root transformation of scores. Therefore, we used square root transformed BDI scores in all analyses that involved the BDI.

In the regression models, we tested for an overall time effect as well as for effects at post-treatment, 6 months, and 12 months. These analyses were performed using PROC GENMOD, SAS version 9.1 software (SAS Institute, Inc., Carey, NC, USA, 2003). Although the primary analyses comparing the two study groups over time included all subjects with at least some post-treatment or follow-up data, regardless of session attendance (intent-to-treat analyses), we repeated the analyses on the subsample who attended all four treatment sessions to examine whether results differed for treatment completers. We also performed multiple imputation analyses for the outcome measures to assess the effect of missing data due to patients who were randomized but provided no follow-up data ($n = 10$ including the two participants who were withdrawn) or incomplete follow-up data ($n = 15$). Baseline and follow-up values of the outcome measures, as well as a variable indicating assigned treatment, were used in the imputation of the missing values, which used a Markov Chain Monte Carlo (MCMC) method assuming an arbitrary missing data pattern and multivariate normality and a single chain to create five imputations using 200 burn-in iterations before the first imputation and 100 iterations between imputations. PROC MI was used to generate the imputed data sets and PROC MIANALYZE was used to combine the results of the GEE analyses of imputed data sets and generate valid statistical inferences (SAS Version 9.1 software programs).

As recommended by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials-II (IMMPACT-II) (Dworkin et al., 2005), we compared the two study groups in terms of clinically meaningful improvement. To examine whether the two groups differed in proportions of participants who showed clinically meaningful improvement on the outcome measures at the primary endpoint (12 months), we used logistic regression analyses. We conservatively considered a change of $\geqslant 50\%$ in pain intensity from baseline to 12 months to be meaningful. This criterion was used to assess response rates in previous pain treatment trials (Serpell and Neuropathic Pain Study Group, 2002) and was found to correspond to patient appraisals of "very much improved" (Farrar et al., 2001). Because clinically significant improvement has not been defined for the activity interference measure, we compared the proportions of the PMT and SCM groups with a score of 0 on this measure at 12 months. For the MFIQ scales, we used logistic regression analyses to compare the PMT and SCM groups in terms of proportion with low impairment at 12 months, adjusting for baseline impairment category. For the BDI, we used logistic regression to compare the PMT and SCM groups in terms of proportion with scores $\geqslant 21$ at 12 months, adjusting for baseline category.

We used analysis of covariance (ANCOVA) to compare the two treatment groups in terms of pre- to post-treatment change in TMD knowledge. We used Mann–Whitney tests to compare the groups on post-treatment ratings of helpfulness of treatment components (because they showed skewed distributions, with more ratings at the high end) and on the mean of the two overall helpfulness and treatment satisfaction ratings ($r = 0.90$ between these ratings).

We did not make a statistical adjustment (e.g., Bonferroni method, which inflates type II error and reduces statistical power) for multiple comparisons. We expected at least moderate correlations of the measures within each domain (outcomes, beliefs, and coping) and thus fairly similar results across the measures within each domain; findings of a significant difference between the two study groups on multiple related measures would support an interpretation that the treatment had an effect on outcomes (or beliefs or coping). We included multiple measures within each domain because it would be of interest to patients, clinicians, and researchers to know whether the CBT intervention had a different impact on, for example, interference versus pain intensity. Similarly, it would be of interest to know if process variables change differentially with CBT. Given these considerations, we specified a priori the primary endpoint, and as recommended for these circumstances (Rothman, 1990; Perneger, 1998; Schulz and Grimes, 2005), did not adjust for multiple comparisons of secondary endpoints.

3. Results

3.1. Sample characteristics, session attendance, and follow-up response

The two study groups did not differ significantly in any sociodemographic characteristic examined (Table 1). The predominance (86%) of females among the participants is consistent with the well-known predominance of women among patients with TMD (Carlsson and LeResche, 1995; Dimitroulis, 1998). The two study groups also did not differ significantly at baseline in chronic pain grade (II high: 29.2%, 30.3%; III: 31.9%, 32.9%; IV: 38.9%, 36.8%; PMT and SCM, respectively; χ^2 test, $P = 0.97$) or duration of current pain episode [PMT median (interquartile range) = 13.5 months (4–78 months), SCM median (interquartile range) = 17.5 months (4–72 months); Mann–Whitney test, $P = 0.99$].

Table 1
Baseline sociodemographic characteristics of study participants

Characteristic	PMT ($n = 72$)	SCM ($n = 76$)	P-value[a]
Age, mean (SD), years	38.9 (11.6)	35.7 (10.9)	0.09
Female, %	86.1	86.8	0.90
Education, %			0.37
High school or less	26.4	17.1	
Some college or vocational/technical	37.5	44.7	
College graduate	36.1	38.2	
White, %	87.5	81.6	0.24
Working full-time, %	44.4	36.8	0.35
Marital status, %			0.91
Married or living as married	48.6	48.7	
Never married	29.2	31.6	
Widowed, separated, or divorced	22.2	19.7	

PMT, Pain Management Training; SCM, Self-Care Management.
[a] From t-test for age and χ^2 tests for other variables.

In the RDC/TMD diagnostic classification system, patients may have more than one diagnosis. Almost all study participants had an RDC/TMD Group I (muscle) diagnosis (PMT 97.2%, SCM 94.7%), with somewhat higher proportions having a Ib (myofascial pain with limited opening) as compared with a Ia (myofascial pain) diagnosis (PMT: 44.4% Ia, 52.8% Ib; SCM: 35.5% Ia, 59.2% Ib; PMT versus SCM comparison $P = 0.45$). Slightly more than half of the participants in each group had a IIIa (arthralgia) diagnosis (PMT: 56.9%, SCM: 55.3%; $P = 0.84$). Among the nine most common RDC/TMD diagnoses in this sample, only two differed significantly between groups. More SCM (22.4%) than PMT (6.9%) participants had a IIa (disc displacement with reduction) diagnosis (χ^2 test, $P = 0.008$). More PMT (9.7%) than SCM (1.3%) participants had a IIIb (osteoarthritis of the TMJ) diagnosis (Fisher's exact test, $P = 0.03$).

Fig. 1 provides information concerning intervention session completion and follow-up assessment completion for all randomized participants. The PMT and SCM groups did not differ significantly in the proportion of randomized participants who completed at least one follow-up assessment (PMT: 94%, SCM: 96%; χ^2 test, $P = 0.49$) or in the number of follow-up assessments completed (χ^2 test, $P = 0.83$). Among all patients randomized ($N = 158$), all four intervention sessions were completed by 57 (74%) PMT participants and 53 (67%) SCM participants (χ^2 test, $P = 0.34$). Nine (12%) PMT participants and seven (9%) SCM participants did not attend any sessions; 11 (14%) PMT and 19 (24%) SCM participants attended one to three sessions. The sample analyzed for this report consisted of the 148 study participants who provided some post-treatment or follow-up data. Among these 148 participants, the proportion of participants who completed all four intervention sessions did not differ between treatment groups (79% of those assigned to PMT and 70% of those assigned to SCM; χ^2 test, $P = 0.19$).

3.2. PMT versus SCM: changes on outcome measures

At baseline, almost all study participants reported that pain interfered with daily activities (Table 2). On average in the sample, pain intensity was moderately high (6.8 in each group) and depressive symptoms were in the mild to moderate range of severity. As has been found in past research with TMD patients (Stegenga et al., 1993b), participants reported more masticatory than non-masticatory jaw use limitations. The two study groups did not differ significantly at baseline on any outcome measure.

Adjusting for baseline scores, across the three follow-up assessments, the PMT group had significantly better outcomes as compared with the SCM group on all outcome measures: pain-related activity interference (the

Table 2
Outcomes for patients in PMT versus SCM

Measure/time	PMT	SCM	*P*-value[a]	Overall *P*-value[b]
Activity interference, % (n) with none				
Baseline	1.4% (1)	5.3% (4)	0.37	
Post	12.9% (9)	5.6% (4)	0.05	
6 months	36.8% (25)	11.1% (8)	0.0002	0.0002
12 months	35.3% (24)	12.9% (9)	0.004	
Characteristic pain intensity, mean (SD)				
Baseline	6.8 (1.7)	6.8 (1.7)	0.97	
Post	5.2 (1.9)	5.2 (2.1)	0.82	
6 months	4.0 (2.5)	4.8 (2.1)	0.03	0.04
12 months	3.9 (2.6)	4.7 (2.3)	0.02	
Beck Depression Inventory,[c] mean (SD)				
Baseline	13.4 (8.6)	13.4 (8.8)	0.99	
Post	8.8 (9.3)	11.0 (10.6)	0.046	
6 months	8.9 (9.8)	10.8 (8.6)	0.37	0.02
12 months	8.3 (9.1)	11.4 (10.1)	0.03	
MFIQ Non-Masticatory, mean (SD)				
Baseline	0.38 (0.19)	0.36 (0.16)	0.44	
Post	0.26 (0.20)	0.29 (0.17)	0.058	
6 months	0.22 (0.19)	0.26 (0.18)	0.04	0.02
12 months	0.20 (0.19)	0.26 (0.16)	0.02	
MFIQ Masticatory, mean (SD)				
Baseline	0.60 (0.26)	0.56 (0.25)	0.31	
Post	0.48 (0.26)	0.54 (0.23)	0.005	
6 months	0.42 (0.26)	0.48 (0.26)	0.009	0.0004
12 months	0.40 (0.27)	0.50 (0.25)	0.0001	

PMT, Pain Management Training; SCM, Self-Care Management; MFIQ, Mandibular Function Impairment Questionnaire.
Mean values shown are observed (not adjusted for baseline differences).

[a] Activity interference was analyzed as none versus some due to the number of patients with scores of 0 at follow-up; for activity interference, *P*-values for baseline comparisons are from χ^2 tests and *P*-values for post and follow-up comparisons are from logistic regression analyses fit using GEE and adjusting for baseline values of activity interference. For all other outcome measures, *P*-values for baseline comparisons are from *t*-tests and for post and follow-up comparisons are from linear regression analyses (fit using GEE) adjusting for baseline values of the outcome variable.

[b] Overall *P*-value indicates the statistical significance of the difference between the two study groups on average over the three follow-up assessments, adjusting for baseline outcome measure scores (from logistic regression for activity interference and linear regression for all other measures).

[c] Because the distribution of Beck Depression Inventory scores was skewed, square root-transformed values were used in the analyses. For ease of interpretation, the mean and SD shown have been transformed back to the original scale.

primary outcome), characteristic pain intensity, depression (Beck Depression Inventory), and the MFIQ Non-Masticatory and Masticatory scales (Table 2). The outcome measures exhibited different patterns of between-group differences across the three follow-up assessments, but all showed statistically significant differences at 12 months. The adjusted (for baseline levels) mean differences (PMT–SCM) and 95% CI at 12 months were −0.92 (−1.70, −0.14) for characteristic pain intensity, −2.93 (−5.35, −0.50) for the BDI, −0.13 (−0.19, −0.07) for MFIQ masticatory limitations, and −0.06 (−0.11, −0.01) for MFIQ non-masticatory jaw limitations. The odds of reporting no activity interference at 12 months, after adjusting for baseline interference, were four times greater in the PMT group than in the SCM group (OR = 4.2; 95% CI = 1.7, 10.2).

In the sensitivity analyses using multiple imputation for missing values, the statistical significance (i.e., whether the *P*-value was < 0.05) of the difference between PMT and SCM did not change for any outcome measure, and the adjusted mean differences between groups were essentially unchanged. In the sensitivity analyses adjusting for RDC/TMD diagnosis, there was a small attenuation of the treatment group differences, but for all outcome measures, the difference between treatment groups remained statistically significant at 12 months. In the analyses of data from participants who completed all four sessions, overall differences between PMT and SCM were greater than in the intent-to-treat analyses for all outcome measures except the BDI and activity interference.

3.3. PMT versus SCM: changes on process measures

As compared with the SCM group, the PMT group showed significantly greater overall improvement and, with only one exception, significantly greater improvement at each follow-up assessment on each belief and catastrophizing measure (Table 3). The two groups did not differ in overall change on three of the four coping scales. However, there was a significant difference ($P < 0.0001$) between groups over time in the use of relaxation techniques to cope with pain. The SCM group showed no change over time in the use of relaxation. An increase in the PMT group at post-treatment was followed by a decrease over the next 9 months, but still greater use at 12 months relative to baseline. Analyses limited to treatment completers yielded comparable results, with no meaningful differences as compared with the intent-to-treat analyses.

3.4. Clinically significant change on outcome measures

Among the patients who provided 12-month data, 50% of the PMT group but only 29% of the SCM group had characteristic pain intensity decreases of $\geqslant 50\%$ from baseline to 12 months (logistic regression, $P = 0.01$). Among treatment completers, the superiority of PMT was even greater: 55% of the PMT group and 24% of the SCM group achieved this level of pain relief (logistic regression, $P = 0.001$).

The superiority of the PMT intervention in terms of clinically significant change was also seen on the measure of masticatory jaw use limitations. At baseline, 17% of the PMT subjects were in the low jaw impairment

Table 3
Process measure changes in PMT versus SCM groups

Measure/time	PMT	SCM	*P*-value[a]	Overall *P*-value[b]
Beliefs				
SOPA Disability, *mean* (*SD*)				
Baseline	2.0 (0.8)	1.8 (0.8)	0.14	
Post	1.4 (0.9)	1.6 (0.8)	0.0004	
6 months	1.2 (0.8)	1.6 (0.8)	<0.0001	<0.0001
12 months	1.3 (1.0)	1.5 (0.8)	0.004	
SOPA Harm, *mean* (*SD*)				
Baseline	2.0 (0.6)	1.8 (0.6)	0.02	
Post	1.4 (0.7)	1.7 (0.7)	0.0002	
6 months	1.4 (0.7)	1.7 (0.6)	0.0001	<0.0001
12 months	1.4 (0.8)	1.6 (0.7)	0.003	
SOPA Control, *mean* (*SD*)				
Baseline	1.8 (0.7)	1.8 (0.7)	0.85	
Post	2.8 (0.7)	2.2 (0.8)	<0.0001	
6 months	2.9 (0.7)	2.2 (0.8)	<0.0001	<0.0001
12 months	2.9 (0.9)	2.2 (0.8)	<0.0001	
Self-efficacy, *mean* (*SD*)				
Baseline	4.8 (2.1)	5.0 (2.1)	0.73	
Post	6.4 (1.9)	5.3 (1.9)	<0.0001	
6 months	6.7 (2.3)	5.6 (2.1)	0.001	<0.0001
12 months	7.1 (2.3)	5.8 (2.0)	<0.0001	
Catastrophizing				
CSQ, % (n) with score of 0				
Baseline	6% (4)	7% (5)	1.0	
Post	20% (14)	8% (6)	0.03	
6 months	27% (18)	7% (5)	0.0007	<0.0001
12 months	29% (20)	4% (3)	<0.0001	
PCS Rumination, *mean* (*SD*)				
Baseline	9.8 (4.1)	10.0 (3.9)	0.75	
Post	6.0 (4.2)	8.2 (3.9)	0.0003	
6 months	6.0 (4.2)	7.4 (3.8)	0.03	0.0003
12 months	5.7 (4.6)	7.3 (3.5)	0.007	
Coping				
CPCI Task Persistence, *mean* (*SD*)				
Baseline	4.4 (1.9)	4.2 (1.6)	0.68	
Post	4.6 (1.8)	4.6 (1.8)	0.91	
6 months	4.4 (2.0)	4.4 (1.8)	0.81	0.84
12 months	4.5 (2.0)	4.5 (1.8)	0.88	
CPCI Coping Self-statements, *mean* (*SD*)				
Baseline	2.7 (1.7)	2.3 (1.6)	0.11	
Post	3.2 (1.8)	2.5 (1.8)	0.05	
6 months	2.6 (1.8)	2.2 (1.8)	0.47	0.30
12 months	2.6 (1.9)	2.3 (1.9)	0.95	
CPCI Relaxation, *mean* (*SD*)				
Baseline	2.2 (1.5)	2.1 (1.4)	0.67	
Post	3.7 (1.6)	2.3 (1.4)	<0.0001	
6 months	3.1 (1.7)	2.1 (1.4)	<0.0001	<0.0001
12 months	3.0 (1.6)	2.1 (1.4)	0.0004	
CPCI Rest, *mean* (*SD*)				
Baseline	3.0 (1.8)	3.3 (1.7)	0.46	
Post	2.8 (1.9)	3.0 (1.7)	0.66	
6 months	2.8 (1.94)	2.5 (1.8)	0.17	0.84
12 months	2.6 (1.9)	2.8 (1.8)	0.61	

PMT, Pain Management Training; SCM, Self-Care Management; SOPA, Survey of Pain Attitudes; CSQ, Coping Strategies Questionnaire; PCS, Pain Catastrophizing Scale; CPCI, Chronic Pain Coping Inventory.
Mean values shown are observed (not adjusted for baseline differences).

[a] *P*-values for baseline comparisons are from *t*-tests for continuous variables and χ^2 tests for categorical variables, *P*-values for post and follow-up comparisons are from linear regression analyses for continuous variables and logistic regression analysis for categorical variables (fit using GEE) adjusting for baseline values of the outcome variable.

[b] Overall *P*-value indicates the statistical significance of the difference between the two study groups on average over the three follow-up assessments, adjusting for baseline outcome measure scores.

category and 51% in the severe category; at 12 months, these proportions were 46% and 22%, respectively. In the SCM group, 16% were low and 46% severe at baseline versus 24% low and 34% severe at 12 months. The difference between proportions of PMT and SCM participants with low versus moderate/severe masticatory impairment at 12 months, adjusting for baseline impairment category, was highly significant in both the intent-to-treat (logistic regression, $P = 0.0006$) and treatment completer analyses ($P = 0.0035$).

On the MFIQ Non-masticatory scale, the PMT group improved from 39% low and 17% severe impairment at baseline to 74% low and 7% severe at 12 months. In the SCM group, there were 37% in the low and 8% in the severe category at baseline, and 64% low and 6% severe at 12 months. The difference between study groups in proportions with low versus moderate/severe non-masticatory impairment at 12 months, adjusting for baseline impairment category, showed a trend towards statistical significance in the intent-to-treat analyses (logistic regression, $P = 0.09$) and statistical significance in the treatment completer analyses (85% of PMT completers in the low category versus 62% of SCM completers; $P = 0.0075$).

At baseline, the PMT and SCM groups did not differ in proportions with BDI scores above the cutoff (scores $\geqslant 21$) for clinically significant levels of depressive symptoms (25% of PMT and 24% of SCM subjects; $P = 0.81$). At 12 months, the proportion in this category increased slightly in the SCM group (26%). In contrast, in the PMT group, the proportion with clinically significant levels of depressive symptoms at 12 months (10%) was less than half that at baseline. The difference between groups at 12 months, adjusting for baseline status, was statistically significant (logistic regression, $P = 0.016$). The odds of having a BDI score <21 at 12 months, after adjusting for baseline BDI score, were almost four times higher in the PMT than in the SCM group (OR = 3.8; 95% CI, 1.2, 12.0). In the results for treatment completers, there was a similar, but no longer statistically significant, pattern of differences between groups.

3.5. Non-specific factors, TMD knowledge, and treatment helpfulness/satisfaction

3.5.1. Treatment credibility and expectations

At baseline (prior to randomization), study participants rated the credibility of the PMT treatment as higher than that of the SCM treatment [mean (SD) = 6.6 (2.0) for PMT, 5.7 (2.0) for SCM on the 0–10 scale; paired t-test for nonzero difference, $P < 0.0001$]. When we repeated the regression analyses examining changes over time on the outcome and process measures, controlling for patient ratings of the credibility of the treatment to which they were assigned, the PMT group remained significantly superior to the SCM group in improvement on every process measure and on every outcome measure except characteristic pain intensity (after controlling for credibility, overall P for the difference between groups = 0.079). When we controlled for both credibility ratings (one for SCM, one for PMT), the treatment group differences became more highly significant on every process and outcome measure including characteristic pain intensity (after controlling for credibility, overall P for the difference between groups in pain intensity = 0.029).

3.5.2. TMD knowledge

The PMT and SCM groups did not differ at baseline in scores on the TMD knowledge measure [PMT mean (SD) = 72 (15), SCM mean (SD) = 74 (13) on the 0–100 scale; t-test, $P = 0.44$]. Both groups showed significant increases in TMD knowledge from baseline to post-treatment [posttreatment PMT mean (SD) = 79 (15), SCM mean (SD) = 82 (12); paired t-test to assess change, $P = 0.004$ in PMT, $P < 0.0001$ in SCM], with no significant difference in change between groups (ANCOVA, $P = 0.16$).

3.5.3. Post-treatment ratings of treatment helpfulness

Participants in both groups on average rated each treatment component assessed as at least moderately helpful, but those assigned to PMT gave higher ratings for all components shared by the treatments except jaw stretching (which was not prescribed for all patients by their dentists). Median (IQR) ratings for PMT and SCM, respectively, were 8.0 (6.0–10.0) and 7.0 (5.0–9.0) for information about TMD (Mann–Whitney test, $P = 0.007$); 9.0 (8.0–10.0) and 8.0 (6.0–10.0) for learning to check and correct oral posture and jaw habits ($P = 0.006$); 9.0 (5.0–10.0) and 7.0 (5.0–9.0) for jaw stretching ($P = 0.16$); and 8.0 (7.0–9.0) and 6.0 (4.0–8.0) for reading material ($P = 0.004$). The mean of the overall helpfulness and treatment satisfaction ratings was significantly higher in the PMT group [median (IQR) = 9.0 (8.0–10.0), 7.0 (5.0–9.0), PMT and SCM, respectively; $P < 0.001$].

4. Discussion

Among patients receiving usual care in a specialty clinic for chronic TMD pain, a brief CBT intervention, as compared with a self-care education/attention control condition, produced statistically and clinically significant improvement in activity interference, pain, depression, and jaw function over the following year. On the primary outcome, activity interference, the proportion of patients with a score of 0 at 12 months was almost three times higher in the CBT group (35%) than in the control condition (13%). That the CBT intervention produced clinically meaningful changes is also supported by

patients' highly favorable evaluations of the treatment's helpfulness (median = 9 on a 0–10 scale).

We previously reported the short-term effects of the CBT versus control condition on daily electronic diary outcome and process measures completed during treatment (Turner et al., 2005). The groups did not differ significantly in rate of within-person change on any daily outcome measure. However, the data suggested a trend towards greater improvement in activity interference and jaw function in the CBT than in the control group in the last two weeks of treatment, possibly indicating a treatment effect that might become more apparent at longer-term follow-ups. The present results substantiate that conjecture.

CBT patients, compared to the control group, also showed significantly greater short- and long-term changes in pain-related beliefs previously demonstrated to play important roles in pain and disability. During the year after treatment, CBT patients had greater decreases in belief they were disabled, belief their pain signaled harm, and catastrophizing, and greater increases in perceived ability to control pain and related problems. At 12 months, the proportion with catastrophizing scores of 0 was seven times higher in the CBT than in the control group (29% versus 4%). These standard validated questionnaire results replicate and extend our previous daily electronic diary findings (Turner et al., 2005) that the CBT group showed greater short-term improvement in pain beliefs and catastrophizing.

In contrast to our hypothesis that CBT would impact all types of pain coping assessed, the only coping strategy that showed a CBT effect was relaxation, which was emphasized much more than the other coping strategies in the CBT protocol. Increases in use of relaxation in the CBT group relative to the control group were seen during treatment on daily electronic measures (Turner et al., 2005) and in the year after treatment on a standard questionnaire. Relaxation techniques may be particularly important for patients with TMD given evidence of a strong association between jaw pain and daytime jaw, face, and head muscle tension (Glaros et al., 2005).

One difference between the daily diary and the longer-term questionnaire results was in cognitive coping. On our diary measure, which assessed multiple cognitive coping strategies (e.g., challenging negative thoughts, viewing pain in a different light), the CBT group showed a much larger increase than did the control group. We assessed only one type of cognitive coping in the year after treatment (coping self-statements), and this measure showed no treatment group difference. It is possible that the CBT intervention changed other types of cognitive coping that were not assessed after treatment. Further research using measures of cognitive and other coping strategies not assessed in this study (as well as different CBT protocols) is needed to more fully understand the effects of CBT on coping.

The benefits of CBT were not due solely to patient education and expectations. CBT and control groups showed comparable increases in TMD knowledge. Similar to previous studies in which patients with chronic pain rated CBT credibility (Bradley et al., 1987; Goossens et al., 2005), patients had only moderate expectations for the two interventions, perhaps because of their previous experiences with pain treatments. Prior to randomization, patients viewed the CBT intervention as more credible than the control intervention. However, even after adjusting for patient ratings of credibility of their assigned treatment, the CBT group showed significantly better outcomes on all measures except pain intensity (for which there was a trend toward a significant difference). This indicates that the benefits of CBT cannot be explained entirely by patient expectations.

The difference in patients' baseline expectations of the two interventions (approximately one point on a 0–10 scale) is of unknown clinical meaningfulness, but it underscores the importance of assessing patients' treatment expectations, as they may influence treatment adherence and outcomes. In future RCTs, efforts should be made to enhance the credibility of the control condition so that it is equal to that of the active treatment. It may be informative to assess treatment credibility both before and after the first treatment session, as credibility could change with treatment exposure.

There exist only a few reports of one-year follow-ups from RCTs of brief CBT for chronic pain. A four-session intervention aimed at reducing fear-avoidance and increasing activity among patients with back pain reduced activity limitations over a two-year follow-up (Von Korff et al., 2005). Two previous studies conducted in the UW Orofacial Pain Clinic compared usual care with versus without brief CBT. A two-session CBT group produced no observed benefits at three months, but significantly greater improvement in pain and a trend towards significantly less activity interference at one-year (Dworkin et al., 1994). Six individual CBT sessions produced significantly greater improvement in pain and a trend toward significantly less interference at four months, but no significant differences at one-year (Dworkin et al., 2002). The inclusion criteria and patient characteristics were similar in the 2002 study and the current study. However, in the current study, the CBT intervention emphasized strategies to maintain gains after treatment, including ways to prevent and manage pain flare-ups. Furthermore, although follow-up telephone calls were made to both CBT and control group patients, in the CBT condition, the psychologist inquired about adherence to treatment recommendations and engaged in problem-solving with patients who reported difficulties. Periodic contact with patients via "booster sessions" has been used in conjunction with interventions for a variety of physical and mental health

problems (Syrjala et al., 1995; Clark et al., 1999; Zijlstra et al., 2005); research is needed to compare brief therapies with and without booster sessions and to determine whether there is a "dose–response" relationship between booster sessions and outcomes.

The 43% enrollment rate is a study limitation, but comparable to rates in previous RCTs of CBT in clinic settings (Dworkin et al., 1994, 2002). Although study participants and nonparticipants did not differ on any variable examined except age, it is possible that participants differed from the larger population of individuals with disabling chronic TMD pain in ways that might have affected the study results. Further research is needed to assess the generalizability of the study findings to other groups of people with TMD pain as well as to individuals with other chronic pain conditions.

Another study limitation is that treatment provider competence and study protocol adherence were not assessed via direct observation (e.g., audiotape review). Intervention differences in treatment provider background might be viewed as a study limitation in that the superiority of CBT may have reflected treatment provider as well as content differences. However, use of a Ph.D.-level psychologist as treatment deliverer in the control condition carries certain risks (e.g., inadvertent use of CBT techniques, negative influence of psychologists' expectations that the control treatment would not be effective) and the use of bachelors' level patient educators represents a more "real-world" approach, as psychologists would be unlikely to conduct solely educational interventions in clinical settings.

Our findings support the cognitive-behavioral model of chronic pain and the added value of CBT in conjunction with specialty dental treatment, and suggest directions for future research. The impact of this brief CBT intervention in improving patient outcomes beyond improvement associated with TMD clinic treatment, education, attention, and natural history indicates the promise of brief CBT for patients with chronic disabling pain. Given that subgroups of patients with chronic pain can be identified reliably based on levels of pain, functioning, and distress, and that subgroup membership predicts response to pain treatments (Von Korff et al., 1992; Keefe et al., 2004), a logical next step would be to tailor brief treatments based on patient subgroup or other characteristics (e.g., depression, anxiety, and sleep disturbance) (Keefe et al., 2004). There is also a need to evaluate stepped care approaches (Von Korff and Moore, 2001) to determine whether patients who fail to improve substantially with brief therapy benefit from further, more intensive, therapy. Finally, many eligible patients declined to participate in our study because of the requirement to attend four sessions; this suggests a need to develop and evaluate telephone- and internet-based CBT interventions. Telephone calls have been used effectively as the sole vehicle of treatment delivery and to supplement in-person CBT interventions (Rene et al., 1992; Mohr et al., 2000; Moore et al., 2000; Holroyd et al., 2001; Damush et al., 2003). Internet-based interventions with and without telephone calls have been demonstrated effective for chronic back pain (Lorig et al., 2002; Buhrman et al., 2004). The relative value of such novel delivery methods versus traditional clinic-based CBT for chronic pain remains to be determined.

Acknowledgments

Funding for this study was provided by the National Institute of Dental and Craniofacial Research Grant P01 DE08773. The authors wish to express deep appreciation to the faculty and staff of the University of Washington Orofacial Pain Clinic, especially Kimberly H. Huggins, RDH, BS, Clinical Research Manager, and Edmond Truelove, DDS, MSD, Professor and Chair of the University of Washington Department of Oral Medicine. The authors also gratefully acknowledge the invaluable assistance of Heather Brister, Eileen Gaspar, Katie Klein, Craig Sawchuk, and Kathy Scott in the conduct of this study. Finally, the authors would like to acknowledge the pioneering contributions of Samuel F. Dworkin, DDS, PhD, in developing biobehavioral treatments for chronic TMD pain and laying the groundwork for this clinical trial.

References

Addis ME, Hatgis C, Bourne L, Krasnow AD, Jacob K, Mansfield A. Effectiveness of cognitive-behavioral treatment for panic disorder versus treatment as usual in a managed care setting. J Consult Clin Psychol 2004;72:625–35.

Astin JA, Beckner W, Soeken K, Hochberg MC, Berman B. Psychological interventions for rheumatoid arthritis: a meta-analysis of randomized controlled trials. Arthritis Rheum 2002;47:291–302.

Beck AT, Beamesderfer A. Assessment of depression: the depression inventory. In: Pichot P, editor. Modern problems in pharmacopsychiatry, Vol. 7. Basel, Switzerland: Karger; 1974. p. 151–69.

Beck AT, Rush AJ, Shaw BF, Emery G. Cognitive therapy of depression. New York: Guilford Press; 1979.

Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. Clin Psychol Rev 1988;8:77–100.

Bernstein DA, Borkovec TD, Hazlett-Stevens H. New directions in progressive relaxation training. Westport, Connecticut: Praeger; 2000. 171 pp.

Borkovec TD, Nau SD. Credibility of analogue therapy rationales. J Behav Ther Exp Psychiatry 1972;3:257–60.

Bradley LA, Young LD, Anderson KO, Turner RA, Agudelo CA, McDaniell LK, et al. Effects of psychological therapy on pain behavior of rheumatoid arthritis patients: treatment outcome and six-month follow-up. Arthritis Rheum 1987;30:1105–14.

Brister H, Turner JA, Aaron LA, Mancl L. Self-efficacy is associated with pain, functioning, and coping among patients with chronic temporomandibular disorder pain. J Orofac Pain, in press.

Buhrman M, Faltenhag S, Strom L, Andersson G. Controlled trial of internet-based treatment with telephone support for chronic back pain. Pain 2004;111:368–77.

Butterworth JC, Deardorff WW. Psychometric profiles of craniomandibular pain patients: identifying specific subgroups. J Craniomandib Practice 1987;5:225–32.

Carlsson GE, LeResche L. Epidemiology of temporomandibular disorders. In: Sessle BJ, Bryant PS, Dionne RA, editors. Temporomandibular disorders and related pain conditions, progress in pain research and management, Vol. 4. Seattle: IASP Press; 1995. p. 211–26.

Chen E, Cole SW, Kato PM. A review of empirically supported psychosocial interventions for pain and adherence outcomes in sickle cell disease. J Pediatr Psychol 2004;29:197–209.

Clark DM, Salkovskis PM, Hackmann A, Wells A, Ludgate J, Gelder M. Brief cognitive therapy for panic disorder: a randomized controlled trial. J Consult Clin Psychol 1999;67:583–9.

Damush TM, Weinberger M, Perkins SM, Rao JK, Tierney WM, Qi R, et al. The long-term effects of a self-management program for inner-city primary care patients with acute low back pain. Arch Intern Med 2003;163:2632–8.

Diggle PJ, Heagerty P, Liang KY, Zeger SL. Analysis of longitudinal data. New York: Oxford University Press; 2002.

Dimitroulis G. Temporomandibular disorders: a clinical update. Br Med J 1998;317:190–4.

Dozois DJA, Dobson KS, Wong M, Hughes D, Long A. Predictive utility of the CSQ in low back pain: individual vs. composite measures. Pain 1996;66:171–80.

Dworkin RH, Turk DC, Farrar JT, Haythornthwaite J, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113: 9–19.

Dworkin SF. Behavioral characteristics of chronic temporomandibular disorders: diagnosis and assessment. In: Sessle BJ, Bryant PS, Dionne RA, editors. Temporomandibular disorders and related conditions, Vol. 4. Seattle: IASP Press; 1995a. p. 175–92.

Dworkin SF. Personal and social impact of orofacial pain. In: Fricton JR, Dubner R, editors. Orofacial pain and temporomandibular disorders. New York: Raven Press; 1995b. p. 15–32.

Dworkin SF, Huggins KH, LeResche L, Von Korff M, Howard J, Truelove E, Sommers E. Epidemiology of signs and symptoms in temporomandibular disorders: clinical signs in cases and controls. J Am Dent Assoc 1990a;120:273–81.

Dworkin SF, Von Korff M, Whitney CW, Le Resche L, Dicker BG, Barlow W. Measurement of characteristic pain intensity in field research. Pain 1990(Suppl. 5):S290.

Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders: review, criteria, examinations and specifications, critique. J. Craniomandib. Disord.: Facial & Oral Pain 1992;6:301–55.

Dworkin SF, LeResche L, Von Korff M. Studying the natural history of TMD: epidemiologic perspectives on physical and psychological findings. In: Vig KD, Vig PS, editors. Clinical research as the basis for clinical practice. Ann Arbor: University of Michigan; 1989. p. 39–60.

Dworkin SF, Turner JA, Mancl L, Wilson L, Massoth D, Huggins KH, et al. A randomized clinical trial of a tailored comprehensive care treatment program for temporomandibular disorders. J Orofac Pain 2002;16:259–76.

Dworkin SF, Turner JA, Wilson L, Massoth D, Whitney C, Huggins KH, et al. Brief group cognitive-behavioral intervention for temporomandibular disorders. Pain 1994;59:175–87.

Eccleston C, Morley S, Williams AC, Yorke L, Mastroyannopoulou K. Systematic review of randomised controlled trials of psychological therapy for chronic pain in children and adolescents, with a subset meta-analysis of pain relief. Pain 2002;99:157–65.

Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001;94:149–58.

Geisser ME, Roth RS, Robinson ME. Assessing depression among persons with chronic pain using the Center for Epidemiological Studies-Depression Scale and the Beck Depression Inventory: a comparative analysis. Clin J Pain 1997;13:163–70.

Glaros AG, Williams K, Lausten L. The role of parafunctions, emotions and stress in predicting facial pain. J Am Dent Assoc 2005;136:451–8.

Gonzalez VM, Stewart A, Ritter PL, Lorig K. Translation and validation of arthritis outcome measures into Spanish. Arthritis Rheum 1995;38:1429–46.

Goossens MEJB, Vlaeyen JWS, Hidding A, Kole-Snijders A, Evers SMAA. Treatment expectancy affects the outcome of cognitive-behavioral interventions in chronic pain. Clin J Pain 2005; 21:18–26.

Holroyd KA, O'Donnell F, Stensland M, Lipchik GL, Cordingley GE, Carlson BW. Management of chronic tension-type headache with tricyclic antidepressant medication, stress management therapy, and their combination. JAMA 2001;285:2208–15.

Jensen MP, Karoly P. Control beliefs, coping efforts, and adjustment to chronic pain. J Consult Clin Psychol 1991;59:431–8.

Jensen MP, Karoly P. Pain-specific beliefs, perceived symptom severity, and adjustment to chronic pain. Clin J Pain 1992;8:123–30.

Jensen MP, Turner JA, Romano JM, Lawler BK. Relationship of pain-specific beliefs to chronic pain adjustment. Pain 1994;57: 301–9.

Jensen MP, Turner JA, Romano JM, Strom SE. The Chronic Pain Coping Inventory: development and preliminary validation. Pain 1995;60:203–16.

Keefe FJ, Brown GK, Wallston KA, Caldwell DS. Coping with rheumatoid arthritis pain: catastrophizing as a maladaptive strategy. Pain 1989;37:51–6.

Keefe FJ, Caldwell DS. Cognitive behavioral control of arthritis pain. Med Clin North Am 1997;81:277–90.

Keefe FJ, Caldwell DS, Queen KT, Gil KM, Martinez S, Crisson JE, et al. Pain coping strategies in osteoarthritis patients. J Consult Clin Psychol 1987;55:208–12.

Keefe FJ, Caldwell DS, Williams DA, Gil KM, Mitchell D, Robertson C, et al. Pain coping skills training in the management of osteoarthritic knee pain: a comparative study. Behav Ther 1990;21:49–62.

Keefe FJ, Rumble ME, Scipio CD, Giordano LA, Perri LM. Psychological aspects of persistent pain: current state of the science. J Pain 2004;5:195–211.

Kendall PC, Hollon SD, Beck AT, Hammen CL, Ingram RE. Issues and recommendations regarding use of the Beck Depression Inventory. Cog Ther Res 1987;11:289–99.

LeResche L. Epidemiology of temporomandibular disorders: implications for the investigation of etiologic factors. Crit Rev Oral Biol Med 1997;8:291–305.

Lorig K, Stewart A, Ritter PL, Gonzalez VM, Laurent D, Lynch J. Outcome measures for health education and other health care interventions. Thousand Oaks, CA: Sage Publications; 1996.

Lorig KR, Laurent DD, Deyo RA, Marnell ME, Minor MA, Ritter PL. Can a back pain e-mail discussion group improve health status and lower health care costs? Arch Intern Med 2002;162:792–6.

Love AW. Depression in chronic low back pain patients: diagnostic efficiency of three self-report questionnaires. J Clin Psychol 1987;43:84–9.

Martin MY, Bradley LA, Alexander RW, Alarcon GS, Triana-Alexander M, Aaron LA, et al. Coping strategies predict disability in patients with primary fibromyalgia. Pain 1996;68:45–53.

Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA 2001;285:1987–91.

Mohr DC, Likosky W, Bertagnolli A, Goodkin DE, Van Der Wende J, Dwyer P, et al. Telephone-administered cognitive-behavioral therapy for the treatment of depressive symptoms in multiple sclerosis. J Consult Clin Psychol 2000;68:356–61.

Moore JE, Von Korff M, Cherkin DC, Saunders K, Lorig K. A randomized trial of a cognitive-behavioral program for enhancing back pain self care in a primary care setting. Pain 2000;88:145–53.

Morley S, de C. Williams AC, Black S. A confirmatory factor analysis of the Beck Depression Inventory in chronic pain. Pain 2002;99:289–98.

Morley S, Eccleston C, Williams A. Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. Pain 1999;80:1–13.

Ohrbach R, Dworkin SF. Five-year outcomes in TMD: relationship of changes in pain to changes in physical and psychological variables. Pain 1998;74:315–26.

Osman A, Barrios FX, Gutierrez PM, Kopper BA, Merrifield T, Grittmann L. The Pain Catastrophizing Scale: further psychometric evaluation with adult samples. J Behav Med 2000;23:351–65.

Perneger TV. What's wrong with Bonferroni adjustments. Br Med J 1998;316:1236–8.

Rene J, Weinberger M, Mazzuca SA, Brandt KD, Katz BP. Reduction of joint pain in patients with knee osteoarthritis who have received monthly telephone calls from lay personnel and whose medical treatment regimens have remained stable. Arthritis Rheum 1992;35:511–5.

Rogers WH, Adler DA, Bungay KM, Wilson IB. Depression screening instruments made good severity measures in a cross-sectional analysis. J Clin Epidemiol 2005;58:370–7.

Rosenstiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and current adjustment. Pain 1983;17:33–44.

Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology 1990;1:43–6.

Rudy TE, Turk DC. Integrating behavioral and dental treatments: utility of customizing protocols. In: Sessle BJ, Bryant PS, Dionne RA, editors. Temporomandibular disorders and related pain conditions, Vol. 4. Seattle: IASP Press; 1995. p. 351–62.

Rudy TE, Turk DC, Zaki HS, Curtin HD. An empirical taxometric alternative to traditional classification of temporomandibular disorders. Pain 1989;36:311–20.

Safren SA, Heimberg RG, Juster HR. Clients' expectancies and their relationship to pretreatment symptomatology and outcome of cognitive-behavioral group treatment for social phobia. J Consult Clin Psychol 1997;65:694–8.

Schulz KF, Grimes DA. Multiplicity in randomised trials i: endpoints and treatments. Lancet 2005;365:1591–5.

Serpell MG, Neuropathic Pain Study Group. Gabapentin in neuropathic pain syndromes: a randomised, double-blind, placebo-controlled trial. Pain 2002;99:557–66.

Shapiro DA, Barkham M, Stiles WB, Hardy GE, Rees A, Reynolds S, Startup M. Time is of the essence: a selective review of the fall and rise of brief therapy research. Psychology and Psychotherapy: Theory, Research, and Practice 2003;76:211–35.

Stegenga B, de Bont LG, Dijkstra PU, Boering G. Short-term outcome of arthroscopic surgery of temporomandibular joint osteoarthrosis and internal derangement: a randomized controlled clinical trial. Br J Oral Maxillofac Surg 1993a;31:3–14.

Stegenga B, de Bont LGM, de Leeuw R, Boering G. Assessment of mandibular function impairment associated with temporomandibular joint osteoarthrosis and internal derangement. J Orofac Pain 1993b;7:183–95.

Strong J, Ashton R, Chant D. The measurement of attitudes towards and beliefs about pain. Pain 1992;48:227–36.

Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. Psychol Assess 1995;7:524–32.

Sullivan MJL, Stanish W, Waite H, Sullivan M, Tripp DA. Catastrophizing, pain, and disability in patients with soft-tissue injuries. Pain 1998;77:253–60.

Suvinen TI, Hanes KR, Gerschman JA, Reade PC. Psychophysical subtypes of temporomandibular disorders. J Orofac Pain 1997;11:200–5.

Syrjala KL. Relaxation and imagery techniques. In: Loeser JD, editor. Bonica's management of pain. Philadelphia: Lippincott Williams & Wilkins; 2001. p. 1779–88.

Syrjala KL, Donaldson GW, Davis MW, Kippes ME, Carr JE. Relaxation and imagery and cognitive-behavioral training reduce pain during cancer treatment: a controlled clinical trial. Pain 1995;63:189–98.

Turner JA, Dworkin SF, Mancl L, Huggins KH, Truelove EL. The roles of beliefs, catastrophizing, and coping in the functioning of patients with temporomandibular disorders. Pain 2001;92:41–51.

Turner JA, Mancl L, Aaron LA. Brief cognitive-behavioral therapy for temporomandibular disorder pain: effects on daily electronic outcome and process measures. Pain 2005;117:377–87.

Turner JA, Romano JM. Self-report screening measures for depression in chronic pain patients. J Clin Psychol 1984;40:909–13.

Turner JA, Romano JM. Cognitive-behavioral therapy for chronic pain. In: Loeser JD, editor. Bonica's management of pain. Philadelphia: Lippincott Williams & Wilkins; 2001. p. 1751–8.

Underwood MR, Barnett AG, Vickers MR. Evaluation of two time-specific back pain outcome measures. Spine 1999;24:1104–12.

Vlaeyen JWS, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000;85:317–32.

Von Korff M. Epidemiological and survey methods: assessment of chronic pain. In: Turk DC, Melzack R, editors. Handbook of pain assessment. New York: The Guilford Press; 2001. p. 603–18.

Von Korff M, Balderson BHK, Saunders K, Miglioretti DL, Lin EHB, Berry S, et al. A trial of an activating intervention for chronic back pain in primary care and physical therapy settings. Pain 2005;113:323–30.

Von Korff M, Dworkin SF, LeResche L, Kruger A. An epidemiologic comparison of pain complaints. Pain 1988;32:173–83.

Von Korff M, Moore JE. Stepped care for back pain: activating approaches for primary care. Ann Intern Med 2001;134:911–7.

Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain 1992;50:133–49.

Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 1993;52:157–68.

Weydert JA, Ball TM, Davis MF. Systematic review of treatments for recurrent abdominal pain. Pediatrics 2003;111:e1–e11.

Zijlstra GAR, van Haastregt JCM, van Eijk JTM, Kempen GIJM. Evaluating an intervention to reduce fear of falling and associated activity restriction in elderly persons: design of a randomised controlled trial [ISRCTN43792817]. BMC Public Health 2005;5:26.