Randomized Effectiveness Study of Four Therapeutic Strategies for TMJ Closed Lock

INTRODUCTION

Temporomandibular disorders (TMD) affect the temporomandibular joint (TMJ), masticatory muscles, and associated structures. Annually, from 1 to 3% of Americans seek professional care for TMD symptoms (Drangsholt et al., 1999). An estimated 2% of people with TMD have jaw locking from a permanently displaced intra-articular disc—that is, disc displacement without reduction with limited opening, or TMJ closed lock (LeResche, 1995). This advanced disorder can cause significant pain and interfere with jaw movement and function.

For most TMD, clinicians generally agree that non-surgical care should precede surgery. However, for TMJ closed lock, surgery has been described as preferable (Wilkes, 1991; Dimitroulis, 2005; Mercuri and Laskin, 1994). In practice, closed-lock treatment varies widely in invasiveness, based mainly on the provider’s experiences.

Studies to date do not support a standard of care for primary or concomitant closed-lock therapy. A review of closed-lock treatment literature found "no distinguishing effects" among orthotics, physical therapy, arthrocentesis, and arthroscopic surgery (Kropmans et al., 1999). Since then, two randomized trials found no difference between selected nonsurgical treatments (Minakuchi et al., 2001), or between arthroscopic surgery and discectomy (Holmlund et al., 2001).

This report describes the first adequately powered randomized study assessing the effectiveness of four treatment strategies for individuals with TMJ closed lock. The strategies, defined by their initial treatment modalities, were: Medical Management, Rehabilitation, Arthroscopy with Post-operative Rehabilitation, and Arthroplasty with Post-operative Rehabilitation. The null hypothesis was that initial treatment with medical management improves pain and jaw function as much and as quickly as initial treatment by rehabilitation with or without surgery.

MATERIALS & METHODS

Participants and Study Design

This randomized effectiveness study was conducted at the University of Minnesota from June, 1992, to June, 2004. Institutional Review Board approval was obtained and renewed annually during recruitment. Participants were recruited from the University’s TMJ and Orofacial Pain Clinic, HealthPartners TMJ Clinics, and the authors’ private practices. Inclusion criteria were: age 18 to 65 yrs; daily pain in affected joints aggravated by jaw movement and function; duplication of pain when the affected joint was examined; magnetic resonance imaging (MRI) diagnosis of stage III or IV closed lock (Wilkes, 1989); limited mouth opening; and at least two years’ availability. Exclusion criteria, assessed by medical history, included: systemic rheumatic disease; generalized joint pain or swelling; pregnancy; concurrent use of steroids, anti-inflammatories, muscle relaxants, or narcotics; major psychiatric disease; any medical contraindication;
drinking more than 3 alcoholic drinks daily; unwillingness to accept study treatments; and prior TMJ surgery. Eligibility was assessed, surgical risks discussed, consent obtained, and enrollment made by one of the authors. The randomized allocation had two strata defined by duration of decreased opening: "Non-chronic" (under 6 months' duration), and "Chronic". Treatment assignment was concealed from participants and care provider(s) in sealed envelopes until the enrollment procedure was completed. The study coordinator then opened the envelope and informed the participant of the group assignment.

Clinical examinations and questionnaires were completed at baseline and three-, six-, 12-, 18-, 24-, and 60-month follow-ups. A single examiner, blinded to treatment assignment, performed all clinical measures. The examiner's reliability for the Craniomandibular Index (CMI) examination was assessed yearly relative to a gold-standard examiner, with inter-examiner intraclass correlation coefficients between 0.88 and 0.93. The examiner had no contact with participants except during planned clinical evaluations. When participants presented for evaluation, a research coordinator administered questionnaires, instructed them not to discuss treatment with the examiner, and placed a thin tape over both pre-auricular areas to conceal the surgical scar's absence or presence. Given these measures, further evaluation of the blind was deemed unnecessary. Masking participants to treatment would have required sham surgery, so the study was single-blind. The APPENDIX provides further details on participants, study design, and the flow of participants through the study.

Treatment Groups and Interventions
Participants were randomized among 4 treatment strategies labeled by their initial treatment modality: Medical Management, Rehabilitation, Arthroscopy, or Arthroplasty. All 4 treatment approaches represent current standard practice.

Medical Management included education on the participant's condition, with optimistic counseling, a self-help program (Wright and Schiffman, 1995), and a six-day regimen of oral methylprednisolone followed by non-steroidal anti-inflammatory drugs (NSAIDs) for 3-6 wks. Muscle relaxants and over-the-counter analgesics were used as needed. One dentist implemented all treatment.

Rehabilitation included treatment from a dentist, physical therapist, and health psychologist (Fricton et al., 1987a). Participants were randomly assigned to one of two treating dentists (ELS, JRF). Treatment was medical management (above) plus an intra-oral orthotic (splint), physical therapy, and cognitive-behavioral therapy. Physical therapy involved joint mobilization, physical therapy modalities, and a home exercise program. Cognitive-behavioral therapy included assessment of oral habits, maladaptive habits, and psychopathology, and two follow-up sessions focused on education, habit reversal, and improvement of compliance and self-efficacy.

Arthroscopy was performed by a randomly assigned oral surgeon (JQS or RBT), with the patient under general anesthesia (Moses and Poker, 1989). The superior joint space was lavaged, intracapsular adhesions underwent lysis, and intracapsular betamethasone was injected. Success was assessed when the mandible was manually moved through excursive movements.

Arthroplasty was an open-joint surgery performed by one oral surgeon (Clyde Wilkes) with patients under general anesthesia (Wilkes, 1991). A disc repositioning procedure (discoplasty) was attempted. If the tissues had degenerated too greatly, disc removal (discectomy) was performed.

The APPENDIX describes each treatment modality further. All care providers followed their assigned participants for 6 mos, with additional follow-up as needed. Medical Management modalities were extended to the other groups; the surgical groups also received post-operative narcotics. Pre-operatively, surgical patients saw the physical therapist, psychologist, and surgeon. Post-operatively, surgical patients received the same treatments by physical therapists and psychologists as the Rehabilitation participants. If surgical patients needed an orthotic, one dentist inserted and monitored it. All but two Arthroscopy and five Arthroplasty patients received an orthotic. There were 15 bilateral and 8 unilateral arthroscopies, and 13 bilateral and 8 unilateral arthroplasties.

Study Measures
The primary outcome measures were the Craniomandibular Index (CMI) and the modified Symptom Severity Index (SSI). The CMI is a reliable, valid instrument for assessing jaw pain and dysfunction (Fricton and Schiffman, 1986, 1987). The SSI is a self-reported pain measure consisting of 5 reliable, valid subscales measuring sensory and affective intensity, frequency, duration, and tolerability of pain, and a non-specific symptom checklist (Fricton, 1990; Fricton and Schiffman, 1987; Fricton et al., 1987b). Clinical TMD studies, including the present study, have used a modified SSI excluding the symptom checklist (Wright et al., 1995, 2000; Shafer et al., 2001; Fricton et al., 2002).

Adjustment variables included baseline CMI and SSI, baseline depression and somatization (Symptom Checklist-90-Revised; Derogatis, 1992), selected demographic variables (age, sex, and education; IMPATH, Fricton et al., 1987b), and family income (Health Status Questionnaire 2.0, Pearson National Computer Systems [NCS], Bloomington, MN, USA).

Treating clinicians assessed treatment compliance through the six-month follow-up, rating it as none or poor (0), adequate (1), or good to excellent (2). Medication use was assessed from data collected at follow-up visits.

Statistical Procedures
The primary analyses were by intention-to-treat; individuals receiving a second treatment modality during follow-up were analyzed according to their original treatment assignments. Ten participants withdrew after randomization, but before receiving therapy (Appendix Fig.). Eight were examined 5 yrs later; the five-year analyses include them. To assess their influence, we included, in a secondary five-year analysis, only participants receiving treatment.

The planned primary analysis was a separate one-way ANOVA at each follow-up time, unadjusted for multiple comparisons. As will become clear, adjustment was unnecessary. We checked whether combining times would produce a group difference using a repeated-measures analysis (linear mixed model, PROC MIXED, SAS Institute, Cary, NC, USA). Another analysis compared treatment groups at each follow-up time, adjusting for variables specified above (SAS GLM procedure). Other planned analyses compared groups by treatment compliance and medication use. Improvement from baseline was tested in each group at each follow-up time by paired t tests.

Assuming that a less invasive strategy would be preferred unless a more invasive treatment strategy would be superior, we planned 3 contrasting treatments—Medical Management vs. each of the other groups—for both the CMI and SSI, giving 6 total comparisons. The key to sample size specification was a
meaningful improvement in jaw function and TMJ pain. Chronic pain studies have estimated a clinically important pain reduction as up to 50% (Turk et al., 1993). Population norms for the CMI and SSI are 0.10 and 0.03, respectively (Schiffman et al., 1990). The study population’s baseline CMI and SSI means were 0.45 and 0.62, respectively. Based on a 40% decrease in TMD-related pain, a clinically meaningful difference at follow-up would be 0.4 x (0.45 - 0.10) = 0.14 for the CMI, and 0.4 x (0.62 - 0.03) = 0.24 for the SSI. The power calculation assumed a one-way ANOVA with per-comparison alpha of 0.05/6 = 0.0083. With one-sided hypothesis tests, 20 individuals per group gave 80% power per-comparison to detect differences of 0.12 and 0.18 in the CMI and SSI, respectively.

RESULTS
Counts of participants randomized to each treatment group, acceptance of their assigned treatment, and attendance at each follow-up visit are presented (Table 1). Comparison of treatment groups at baseline showed no differences (p > 0.09, Table 2). According to the same baseline characteristics, the ten participants declining their assigned treatment did not differ from those accepting it (p > 0.2).

CMI results (Table 3) and SSI results (Table 4) show 2 rows of p-values. The upper row indicates differences between groups at each follow-up time, unadjusted for multiple comparisons. Group means did not differ at any follow-up for either CMI (p > 0.3) or SSI (p > 0.08); neither was there a between-group difference after the exclusion of participants who declined assigned treatment (p > 0.5; "60-month Treatment Received" in Tables 3 and 4). Repeated-measures analysis of CMI showed no difference between groups in average or time-course. Repeated-measures analysis of SSI showed group differences only if follow-up time was treated as a continuous predictor, i.e., based on the assumption that SSI declined linearly. Then, the groups differed in time trend (p = 0.03): Arthroplasty achieved full effect by 3 mos, while the other groups improved throughout the 60-month follow-up. Adjusted analyses of CMI showed no group differences at any

Table 1. Total Participants Randomized, Total Participants Accepting Randomized Treatment at Baseline, and Total Participants Seen at Six Follow-up Intervals (Distributions by Treatment Group)

<table>
<thead>
<tr>
<th>Treatment Groups</th>
<th>Randomized</th>
<th>Treatment Received</th>
<th>3 Mos</th>
<th>6 Mos</th>
<th>12 Mos</th>
<th>18 Mos</th>
<th>24 Mos</th>
<th>60 Mos, Randomized</th>
<th>60 Mos, Treatment Received</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical Management</td>
<td>29</td>
<td>29</td>
<td>28</td>
<td>28</td>
<td>29</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>Rehabilitation</td>
<td>25</td>
<td>23</td>
<td>22</td>
<td>20</td>
<td>22</td>
<td>21</td>
<td>21</td>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td>Arthroscopic Surgery</td>
<td>26</td>
<td>23</td>
<td>23</td>
<td>21</td>
<td>23</td>
<td>20</td>
<td>22</td>
<td>24</td>
<td>22</td>
</tr>
<tr>
<td>Arthroplasty Surgery</td>
<td>26</td>
<td>21</td>
<td>20</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>20</td>
<td>24</td>
<td>20</td>
</tr>
<tr>
<td>Totals</td>
<td>106</td>
<td>96</td>
<td>93</td>
<td>88</td>
<td>94</td>
<td>88</td>
<td>89</td>
<td>98</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 2. Baseline Measures and Adjustment Variables [Means and (Standard Error of the Mean) by Treatment Group]

<table>
<thead>
<tr>
<th>Treatment Groups</th>
<th>Birth Date (a)</th>
<th>Age (b)</th>
<th>% Females</th>
<th>Education (c)</th>
<th>Family Income (d)</th>
<th>Marital Status (e)</th>
<th>Depression</th>
<th>Somatization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical Management (n = 29)</td>
<td>20,473 (645)</td>
<td>33.7 (1.8)</td>
<td>90</td>
<td>H - 4/29</td>
<td>6.39 (0.45)</td>
<td>S - 4/29</td>
<td>49.6 (2.0)</td>
<td>50.8 (1.7)</td>
</tr>
<tr>
<td>Rehabilitation (n = 25)</td>
<td>21,916 (611)</td>
<td>30.0 (1.7)</td>
<td>100</td>
<td>H - 4/25</td>
<td>5.72 (0.54)</td>
<td>S - 8/25</td>
<td>49.1 (2.4)</td>
<td>50.3 (2.0)</td>
</tr>
<tr>
<td>Arthroscopic Surgery (n = 26)</td>
<td>21,277 (628)</td>
<td>31.8 (1.7)</td>
<td>85</td>
<td>H - 6/26</td>
<td>5.38 (0.39)</td>
<td>S - 6/26</td>
<td>51.0 (1.8)</td>
<td>52.2 (2.0)</td>
</tr>
<tr>
<td>Arthroplasty Surgery (n = 26)</td>
<td>21,264 (646)</td>
<td>31.4 (1.9)</td>
<td>96</td>
<td>H - 10/26</td>
<td>5.38 (0.43)</td>
<td>S - 7/26</td>
<td>52.0 (2.0)</td>
<td>53.5 (1.5)</td>
</tr>
</tbody>
</table>

P-value for difference p = 0.4 p = 0.5 p = 0.2 p = 0.3 p = 0.09 p = 0.3 p = 0.7 p = 0.5

(a) Number of days since 1960 to day of birth.
(b) Age in yrs at start of the study.
(c) Overall, 23% were high school graduates (H), 67% had a Vocational Tech or college education (C), and 10% had a Graduate College education (G).
(d) The response categories were 1 (0 to $10,000) to 9 ($90,000 or more). The highest value, 6.39, represents an average annual family income between $60,000 and $70,000.
(e) Overall, 24% were single (S), 63% were married (M), and 13% were divorced (D).
follow-up (p ≥ 0.09). In adjusted analyses of SSI, Arthroplasty was superior to Medical Management at the six-month follow-up (p = 0.02, with multiple comparisons disregarded); otherwise, groups did not differ. All groups improved over baseline (p < 0.0001) at all follow-up times for both CMI and SSI (lower row of p-values in Tables 3 and 4).

Treatment compliance was inversely associated (p < 0.05) with SSI up to 18 mos but not later. The groups did not differ in compliance (p > 0.09); average compliance scores ranged from 1.1 for Rehabilitation to 1.7 for Medical Management. Adjustment for compliance did not produce group differences for CMI or SSI (p > 0.2). Groups differed in prescription medication use at 3 mos (p = 0.001), but not later. At 3 mos, proportions of participants requiring a prescription analgesic more than once weekly were 17/28 (61%), 2/21 (10%), 9/22 (23%), and 1/19 (5%) for Medical Management, Rehabilitation, Arthroscopy, and Arthroplasty, respectively. Physical therapy visits averaged nine, seven, 12, and 16, respectively.

Each participant received the assigned initial treatment modality for at least 3 mos, after which some participants in each group needed additional interventions for persistent pain and reduced range of motion. Of 29 Medical Management participants, 12 received rehabilitation after 3 mos, and one received arthroscopy after 12 mos (total, 45% receiving a second modality). One Rehabilitation participant received arthroplasty after 6 mos. Two Arthroscopy participants and one Arthroplasty participant received repeat surgeries. One Arthroplasty participant experienced nerve injury that resolved completely.

**DISCUSSION**

The four treatment strategies did not differ in magnitude or timing of improved function or pain relief. Fifty-five percent of those initially treated with medical management alone had symptom resolution, consistent with natural history studies (Sato et al., 1997; Kurita et al., 1998). Just two of 52 Medical Management and Rehabilitation participants received surgery for persistent pain and restricted function.

An intervention study's generalizability derives from its treatments and its population. This effectiveness study compared treatment strategies defined by initial modality. The Medical Management group was not intended to simulate treatment with medical management alone; rather, it simulated a strategy beginning with medical management and using more invasive methods only if required. The study strategies matched clinical practice and are therefore generalizable.

Regarding the study population, the principal symptoms of closed lock are pain and limited opening. By these symptoms, our study population is comparable with closed-lock surgical groups in four other RCTs (Stegenga et al., 1993; Fridrich et al., 1996; Goudot et al., 2000; Holmlund et al., 2001). When the SSI’s pain intensity scale was transformed to 0 to 10, mean baseline pain for this study’s sample was 6.3, and mean maximum interincisal opening was 29.1 mm, compared with ranges of 5.6 to 7.1 and 27.6 to 30 mm in the other four studies.

Criteria for the closed-lock diagnosis included clinical and medical symptoms of pain and limited opening.
MRI evaluation, both having acceptable validity (Schiffman et al., 1989; Liedberg et al., 1996). For Arthroscopy and Arthroplasty participants, all diagnoses were confirmed during surgery. We did not evaluate arthrocentesis, which was used infrequently when this study began. Two randomized trials have compared arthrocentesis and arthroscopic surgery for closed lock (Fridrich et al., 1996; Goudot et al., 2000). Pain reduction was similar for the two treatments, but the latter study found arthroscopic surgery superior for improving mouth opening. Finally, the present study cannot determine how individual modalities (self-care, medications, physical therapy, cognitive-behavioral therapy, splints, or surgery) contributed to a group’s results.

This study provides a basis for advising closed-lock patients that, on average, short-term improvement with regard to pain and function, as measured at 3 mos, is similar for all four treatment strategies. Primary treatment for patients with closed lock should consist of medical management or rehabilitation. Within the context of this study, we were unable to detect any net benefit associated with surgery over that of medical management or rehabilitation at any follow-up period.

ACKNOWLEDGMENTS

We acknowledge Dr. Clyde Wilkes and, from the University of Minnesota, Pat Lenton, Amanda Jackson, Constance Bromaghim, Bruce Lindgren, and Drs. Gary Anderson, Jana Reinhart, and Greg Pelletier. This study was supported by NIH/NIDCR grants R29DE08668, P30-DE09737, N01-DE22635, and R01 DE13421-01, by HealthPartners of Minnesota, Allina, and by the University of Minnesota School of Dentistry.

REFERENCES


Fricton JR, Look JO, Schiffman E, Swift J (2002). Long-term study of temporomandibular joint surgery with alloplastic implants compared with nonimplant surgery and nonsurgical rehabilitation

<table>
<thead>
<tr>
<th>Treatment Groups</th>
<th>Baseline before Treatment, Mean Scores</th>
<th>3 Mos Post-trt.</th>
<th>6 Mos Post-trt.</th>
<th>12 Mos Post-trt.</th>
<th>18 Mos Post-trt.</th>
<th>24 Mos Post-trt.</th>
<th>60 Mos Treatment Intention-to-Treat</th>
<th>60 Mos Treatment Received</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical Management (MM)</td>
<td>n = 29</td>
<td>0.61 (0.23)</td>
<td>0.33 (0.22)</td>
<td>0.31 (0.20)</td>
<td>0.28 (0.26)</td>
<td>0.27 (0.21)</td>
<td>0.21 (0.24)</td>
<td>0.23 (0.25)</td>
</tr>
<tr>
<td>Rehabilitation (NS)</td>
<td>n = 24</td>
<td>0.72 (0.17)</td>
<td>0.42 (0.27)</td>
<td>0.34 (0.22)</td>
<td>0.31 (0.24)</td>
<td>0.29 (0.28)</td>
<td>0.27 (0.23)</td>
<td>0.23 (0.23)</td>
</tr>
<tr>
<td>Arthroscopic Surgery (AS)</td>
<td>n = 26</td>
<td>0.70 (0.19)</td>
<td>0.34 (0.22)</td>
<td>0.34 (0.22)</td>
<td>0.30 (0.21)</td>
<td>0.28 (0.22)</td>
<td>0.25 (0.20)</td>
<td>0.26 (0.20)</td>
</tr>
<tr>
<td>Arthroplasty Surgery (AR)</td>
<td>n = 26</td>
<td>0.76 (0.22)</td>
<td>0.26 (0.24)</td>
<td>0.19 (0.19)</td>
<td>0.20 (0.22)</td>
<td>0.26 (0.24)</td>
<td>0.25 (0.24)</td>
<td>0.28 (0.25)</td>
</tr>
</tbody>
</table>

P-value for between-group difference: p = 0.06 | p = 0.23 | p = 0.083 | p = 0.37 | p = 0.97 | p = 0.86 | p = 0.83 | p = 0.50 |

P-value for within-group improvement from baseline: NA | p < 0.0001 all groups | p < 0.0001 all groups | p < 0.0001 all groups | p < 0.0001 all groups | p < 0.0001 all groups | p < 0.0001 all groups | p < 0.0001 all groups |


