Reduced Neuropsychologic Measures as Predictors of Treatment Outcome in Patients with Temporomandibular Disorders

Marcio L. Grossi, CD, MSc, PhD
PUCRS Dental School
Catholic University of Rio Grande do Sol
Brazil

Michael B. Goldberg, DDS, Dip Perio, MSc
Faculty of Dentistry
University of Toronto

Wasser Pain Management Centre
Department of Dentistry
Mount Sinai Hospital

Toronto, Ontario, Canada

David Locker, BDS, PhD
Faculty of Dentistry
University of Toronto

Howard C. Tenenbaum, DDS, D Perio, PhD, FRCDC
Faculty of Dentistry
CIHR Group in Periodontal Physiology
University of Toronto

Wasser Pain Management Centre
Department of Dentistry
Mount Sinai Hospital

Toronto, Ontario, Canada

Correspondence to:
Dr Howard C. Tenenbaum
Faculty of Dentistry
124 Edward Street
Toronto, Ontario
M5G 1X5 Canada
Fax: (416) 586-8844
E-mail: howard.tenenbaum@utoronto.ca

Aims: *To determine via a prospective investigation whether the presence of neuropsychologic or cognitive deficiencies could be identified in patients with temporomandibular disorders (TMD) and used to predict treatment outcome. This was based on the theory that measurable reductions in neuropsychologic and cognitive function might have a negative impact on treatment outcome in patients with essentially nontraumatic TMD, as has been shown for patients with posttraumatic TMD.* **Methods:** *Various neuropsychologic, psychosocial, and clinical parameters (including but not limited to the Peterson-Peterson Consonant Trigram Test and the California Verbal Learning Test) were used to pretest patients suffering from TMD prior to treatment. Patients were then entered into treatment, after which determination of treatment success was made both by the use of visual analog scales for pain and global transitional outcome measures (eg, "better," responders versus "same/worse," nonresponders). After determination of treatment success was made, treatment response was correlated with the various clinical, cognitive, and neuropsychologic test scores.* **Results:** *Overall, the nonresponders did worse in both the neuropsychologic and psychosocial assessments, with greater memory deficits, sleep disturbances, depression, and fatigue and lower energy levels as compared to responders. Among the best predictors of treatment outcome were the Peterson-Peterson Consonant Trigram Test scores, as well as the scores on the California Verbal Learning Test (ie, poorer test outcomes predicted nonresponse). Neither responders nor nonresponders could be distinguished from one another based on clinical parameters of maximum interincisal opening or muscle tenderness. Three psychosocial variables were also found to be predictors of poor outcome: sleep disturbance, fatigue, and income. Pretreatment pain on chewing was also found to be a reliable predictor of poor treatment outcome.* **Conclusion:** *We conclude that various neuropsychologic, psychosocial, and some clinical parameters may provide pretreatment prediction of treatment outcome in an idiopathic TMD population.*
J OROFAC PAIN 2001;15:329–339.

Key words: temporomandibular disorders, psychological tests, treatment outcome, cognition, prospective study

Temporomandibular disorders (TMD) are characterized largely by facial pain, which is often exacerbated by jaw movement. The most prevalent orofacial pain conditions are musculoskeletal in origin, and of these, TMD are overwhelmingly the most common.[1–4] Despite the high rate of treatment success,[5–7] it appears that 2% to 30% of patients do not improve and may in fact be nonresponsive to therapy, irrespective of the treat-

ment modality used.[5] Usually, this failure has been attributed to a combination of behavioral and psychosocial factors and their interactions with pathophysiologic factors associated with TMD.[8] To date, there is somewhat limited information about those patients who are nonresponsive to treatment because, understandably, most investigations have focused on the development of successful symptom management or characterization of the TMD population.[5,9,10] However, a recent trend in the literature has been to identify and determine the prevalence of psychosocial factors that may be predictive of a TMD patient's response to treatment.[9–14] These include sleep disturbance, decreased energy level, and reduced appetite, as well as problems with memory and concentration.

Neuropsychologic testing (eg, attention, learning and memory, motor skills, verbal and nonverbal skills, comprehension, and expression of language) has been used in psychiatric and nonpsychiatric populations for objective evaluation of performance for many years.[15,16] These tests have also been used to assess the need for cognitive rehabilitation, predict the course of psychiatric and nonpsychiatric illnesses, and reduce diagnostic heterogeneity within disorders. A significantly increased prevalence of apparent neuropsychologic deficits (ie, memory, attention, reaction time deficits) was found in patients whose TMD arose following a motor vehicle accident (MVA), in comparison to patients with nontraumatic or idiopathic TMD.[17] It has also been demonstrated that there are more nonresponders in this posttraumatic group.[18] Accordingly, it was suggested that neuropsychologic deficits may either play an integral role in mediating poor treatment outcome or at least may be predictors of poor treatment response in the posttraumatic TMD population.

In light of the foregoing, the primary objective of this study was to determine the clinical utility of neuropsychologic tests as predictors of treatment outcome in subjects with idiopathic TMD (ie, no history of trauma). This might permit assessment of TMD patients prior to management to predict those patients in whom good outcomes might be expected (responding TMD patients [rTMD]) versus those in whom a poor treatment outcome might occur (nonresponding TMD patients [nrTMD]). Our secondary objective was to determine whether the traditional signs and symptoms of TMD, as well as other relevant psychosocial factors (ie, depression, fatigue, energy level, sleep, educational level), could also be used as predictors of treatment outcome in the population.

Materials and Methods

Population

Patients participating in this study were chosen from those newly diagnosed and seeking treatment for a TMD in the Craniofacial Pain Research Unit of the Wasser Pain Management Centre at Mount Sinai Hospital (Toronto, Canada) and the Facial Pain Clinic at the University of Toronto, Faculty of Dentistry (Toronto, Canada). The Human Ethics Committee of the University of Toronto approved the study and consent form. A stipend was paid to each participant. To be considered for inclusion in this study, subjects had to fulfill the following criteria:

1. Women between the ages of 15 and 45 years
2. Chief complaint of pain (at least 4 times/week for at least 4 weeks) in the temporomandibular joint (TMJ) and/or masticatory muscle region
3. Tenderness to palpation of at least 3 sites in the masticatory muscles and/or the TMJ region and/or limitation in mandibular movement (interincisal opening less than 40 mm).[11,17,19]

Patients were excluded from consideration for the study if their pain was the result of an arthritic condition (osteoarthritis/rheumatoid arthritis) or if their primary pain complaint came about in association with a traumatic injury. In addition, patients who had been or were currently under treatment for a TMD were excluded from the study group. As well, those with metabolic disorders (diabetes, hyperthyroidism); neurologic disorders; vascular disease (migraine, hypertension); neoplasia; or a history of psychiatric/drug abuse conditions were not considered for the study. Given the nature of the neuropsychologic testing format, those patients who reported significant visual, auditory, and/or motor impairments were also excluded from consideration.

Pain Scale

Pain intensity was measured with a 100-mm visual analog scale (VAS),[20,21] based on the subject's own perception of jaw pain at rest and while chewing within the past month. Anchors to both scales were labeled as "no pain" and "extremely severe pain." No other constraints were placed on the subjects.

Neuropsychologic Testing

Previous studies investigating the effects of closed-head injury indicated that these patients could be

differentiated from a group of "control" subjects with no signs of injury, based on the outcome of various neuropsychologic tests. This fact was demonstrated in a TMD population when Goldberg et al[17] showed that certain neuropsychologic tests could be used to differentiate a posttraumatic TMD population from a nontraumatic TMD population. Given that many of the characteristics of nonresponsive TMD or refractory TMD are similar in both the posttraumatic and nontraumatic populations, it was hypothesized that these neuropsychologic tests may also be able to differentiate and predict those who may respond to therapy and those who would be more refractory to treatment. Therefore, similar neuropsychologic assessments to those used by Goldberg et al[17] were used here.

Simple and Complex Multiple-Choice Reaction-Time Test.[22] A computer-based reaction time test was administered to both study groups. Reaction to a simple stimulus (SRT) was recorded based on the speed at which the subject pressed a button held in their dominant hand in response to a circle, square, triangle, or cross. Subjects were provided with either a target or nontarget stimulus (simple-choice reaction-time test) to increase task complexity. Color and internal structure were further added to increase the complexity of each target and nontarget stimulus. Reaction to the target/nontarget stimuli was calculated, along with an assessment of the number of errors made by each subject when they pressed an incorrect button.

California Verbal Learning Test.[23] This test is a 16-item, 4-category "shopping list" that can be used to assess a subject's immediate, short-term, and long-term memory capacity. This test also assesses a subject's ability to categorize lists, thereby probing the memory strategies of the individual. Therefore, variables such as semantic clustering, perseveration, intrusions, and interference with short-term and long-term recall were tested.

Peterson-Peterson Consonant Trigram (CCC).[24] This test assesses immediate memory recall during an "interference." Subjects were asked to repeat 3 consonants presented after being challenged with a continuous mathematical subtraction problem for 3, 9, or 18 seconds. The total number of correct consonants repeated were scored, regardless of the order in which the subject repeated them.

Psychosocial Assessment

To assess the magnitude of depression as a possible predictor of treatment outcome,[9] the long form of the Beck Depression Inventory (BDI) was used.[25] Measures of fatigue and energy level were assessed by a VAS. Sleep patterns may also correlate with treatment outcome[9]; to assess this, a validated 19-item self-administered questionnaire was used (University of Toronto Sleep Assessment Questionnaire [SAQ]). Normative data, test description, reproducibility, and validity of the SAQ have been published elsewhere.[26]

Clinical Examination

Following the neuropsychologic/psychosocial testing component, subjects underwent a complete extraoral and intraoral clinical examination. With the Research Diagnostic Criteria (RDC) as a guide,[11] the extraoral examination included palpation of the masseter, temporalis, and sternocleidomastoid muscles, as well as palpation of the TMJ. Intraorally, the medial pterygoid muscle, lateral pterygoid muscle (or region), and the insertion of the temporalis muscle at the coronoid process were palpated. Although the medial pterygoid is not part of the standard RDC, the experience of the group treating these subjects suggested that this muscle was a useful site to palpate. Scores were assigned by the use of a scoring system (scale of 0 to 3) based on the patients' evoked response to palpation, as described previously.[17] The intraoral examination included a complete dental/periodontal examination to rule out pain of dentoalveolar origin. Maximum unassisted interincisal opening, as well as the level of pain exacerbation following the examination, was measured. In addition, pain intensity at rest and while chewing was assessed. A single examiner performed all examinations. Although this was done in an attempt to reduce variability, we recognize that an intraexaminer variability analysis (Kappa analysis) was not performed. This was not done, so as to reduce the burden that such an examination would have placed on the test subjects strictly for statistical purposes. The validity and reliability of such an examination have been described elsewhere.[11] It should also be pointed out that there were 4 treating clinicians, but as shown in the Results, treatment outcomes did not vary between them, so this was not considered to be a significant confounding factor.

Treatment of TMD

Following the initial examination, patients entered into the treatment phase of the study. All modalities of therapy were provided at the discretion of the treating clinician. However, in each case, a

conservative and reversible treatment approach was the primary consideration in the selection of therapy. Treatment modalities included any or all of the following:

1. Mandibular hard acrylic flat bite plane with full posterior coverage and cingulum coverage in the anterior
2. Low-dose muscle relaxant (cyclobenzaprine, 5 to 10 mg at bedtime for 30 days)
3. Nonsteroidal anti-inflammatory (diflunisal, 500 mg twice daily for 30 days)
4. Physical therapy (moist heat, massage, ultrasound, manipulation).

Each subject's response to treatment was measured following a 6-month course of therapy. This was done through a mailed follow-up questionnaire. Pain intensity at rest and pain intensity while chewing were measured on a VAS. The format was identical to those used for assessment at baseline. For a patient to be considered "improved," a 30% reduction in their baseline VAS had to be recorded. Previous studies have indicated that a 30% reduction in VAS measurements will take into consideration both the inherent variability of the scale as well as issues surrounding "pain memory."[5,20] In addition, improvement was also measured by the use of a global transition judgment, in which subjects were asked to rate themselves as "better," "same," or "worse." These 3 outcomes were reduced to 2 groups—those responding to treatment (better) and those not responding to treatment (same/worse).

Sample Size Calculation/Statistics

Calculation of sample size for 2 independent means as described by Taylor was used.[27] The sample size was calculated from the means and standard deviations based on results of the SRT generated in preliminary studies, which was based on the best available data for this population.[17] This test was chosen because it showed the smallest significant difference between the 2 groups among all reaction-time tests ($P < .05$). It was estimated that 57 patients had to be screened in the initial TMD group to obtain a sample of 17 individuals with nrTMD. The number was increased to 60 in the TMD group and 20 for nrTMD patients to compensate for dropouts.[19]

Given the distribution of data, the Mann-Whitney *U*/Wilcoxon rank sum test was used for continuous variables. Chi-square and Fisher exact tests were used to identify statistically significant differences ($P < .05$) between the 2 groups where categorical data were used. As well, relative risk (RR) and 95% confidence intervals were used to determine the strength of the associations found. For continuous variables, cutoff points were based either on the normative data of the tests, or, when not possible due to a lack of published norms, were set at ± 2 standard deviations from the pain-free population mean.[16] When a cutoff point was selected, both sensitivity and specificity were calculated. Positive test results were indicated as 1, and negative test results were indicated as 0. The nrTMD group was designated as 1, meaning "disease positive," and the rTMD group was designated as 0, meaning "disease negative" for calculation of sensitivity and specificity. The role of confounders was assessed by the use of logistic regression.

Results

Study Population

Only 19% of the new patients presenting to the clinic met the inclusion criteria, and of these, 50% agreed to participate in the study. The most common diagnosis for those presenting to the clinic was myogenous pain with or without TMJ disc-interference disorder. Other common diagnoses for patients screened but not entered into the study included posttraumatic TMD following an MVA (10%), neurologic disorders (9%), headaches (8%), TMJ pain without muscle involvement (7%), atypical facial pain (7%), strictly psychologic disorders (7%), and oral mucosal diseases (6%). Of those initially assessed with the battery of neuropsychologic tests, the follow-up rate was 100%. The majority of the missing values for the 45 variables tested did not reach 5%, and in those that did, no statistically significant difference was found in the distribution of those variables tested across the groups.

The sociodemographic variables in the study group were also assessed. The majority of the population had a high level of education (postsecondary diploma/certificate or higher, 64.4%) and were employed (67.8%). However, the majority of the population also belonged to a lower income group (CDN $39,000 a year or less, 57.9%). The mean age for the test population was 28.3 ± 9.0 years.

Table 1 Neuropsychologic and Psychosocial Test Results (Mean and SD) in Responding and Nonresponding TMD Patients

Dependent variables	Responding TMD group (n = 36)	Nonresponding TMD group (n = 24)	Mann-Whitney *U*/Wilcoxon rank sum test
Simple reaction time (msec)	249 (60)	261 (67)	$P = .3$; NS
Multiple-choice reaction time (msec)	437 (61)	477 (92)	$P = .1$; NS
Multiple-choice reaction time with conflict (msec)	484 (73)	528 (107)	$P = .1$; NS
Multiple-choice reaction time with constraint (msec)	447 (66)	480 (99)	$P = .02$; NS
CVLT-CR (scores 0–80)	60 (8.5)	53 (10)	$P = .005^{\dagger}$
CVLT-CL (scores 0–60)	27 (11)	18 (7.4)	$P = .001^{\ddagger}$
CVLT-P (scores 0–40)	5.3 (6.2)	5.6 (4.2)	$P = .8$; NS
CVLT-I (scores 0–10)	0.6 (1.0)	0.7 (1.4)	$P = .7$; NS
CCC (scores 0–45)	34 (6.3)	30 (6.3)	$P = .006^{\dagger}$
SAQ (scores 0–68)	20 (6.2)	24 (6.8)	$P = .02^{*}$
BDI (scores 0–63)	7.8 (6.4)	11.0 (6.9)	$P = .08$; NS
Fatigue (VAS, 0–100 mm)	46 (27)	68 (25)	$P = .004^{\dagger}$
Energy level (VAS, 0–100 mm)	50 (24)	44 (26)	$P = .4$; NS

$^{*}P < .05$; $^{\dagger}P < .01$; $^{\ddagger}P < .001$.
CVLT = California Verbal Learning Test; CVLT-CR = CVLT immediate recall; CVLT-CL = CVLT semantic clustering; CVLT-P = CVLT = Perseveration; CVLT-I = CVLT = Intrusions; CCC = Peterson-Peterson Consonant Trigram Test; SAQ = Sleep Assessment Questionnaire; BDI = Beck Depression Inventory.

Treatment Outcomes

There was no difference in pretreatment pain at rest as scored on the 100-mm VAS when a post hoc comparison between the rTMD (64 mm) versus nrTMD (63 mm) patients was made. Following a 6-month course of therapy, 60% of patients reported improvement. According to findings obtained with the VAS, pain at rest posttreatment was lower in patients who responded to therapy (rTMD = 21 mm) than in those who did not respond (nrTMD = 60 mm, $P < .001$). The overall percentage agreement between improvement measure using VAS versus improvement measure using global transition judgment was 90%, and the kappa index controlling for observer agreement was 0.78. This level of agreement is considered to be substantial (ie, between 0.61 and 0.80).[28] Once patients were divided into the responding and nonresponding groups, further analysis of each population's pain perceptions could be completed. On the basis of the VAS measurements, the average reduction in pain as compared to baseline pain for rTMD patients was very high (67.1%; $P < .001$), whereas there was virtually no change in VAS measurements for pain in the nrTMD patients as compared to baseline values. There was, however, a significant difference in pretreatment pain on chewing, as shown in VAS measurements, indicating that the responding population had markedly less pain on chewing as compared to the nonresponding population (rTMD = 39 mm; nrTMD = 60 mm; $P < .01$). Consequently, this parameter was also demonstrated to be a predictor of treatment outcome. A similar trend was also shown for pain on chewing posttreatment ($P < .001$). However, as compared to pain at rest, the degree of improvement over baseline was less pronounced (35%) but still significant ($P < .01$).

Neuropsychologic Tests

There was no significant difference in reaction times for either simple or more complex stimuli in the responding and nonresponding subjects, as shown in Table 1. Although a trend toward slower reaction times was evident in the nonresponding population, this difference did not reach statistical significance, and in fact, differences of less than 100 msec are considered to be clinically unimportant in any event.[15,22,29]

On the other hand, rTMD and nrTMD patients could be differentiated on the basis of scores obtained from the California Verbal Learning Test to assess immediate recall (CVLT-CR) and semantic clustering (CVLT-CL). Out of a possible score of 80, the mean score for the rTMD group was 60 ± 8.5, while the nrTMD population scored 53 ±

Table 2 Confounders in Responding and Nonresponding TMD Patients

Independent variables (unit or category)	Responding TMD group (n = 36)	Nonresponding TMD group (n = 24)	Relative risk (RR) (95% CI)*	Significance
Educational level				
Postsecondary diploma/ certificate or higher = 0	74.3	50.0	1.8 (0.9–3.2)	$P = .05$ (NS; chi-square test)
Some education after high school or less = 1	25.7	50.0		
Employment (%)				
Employed = 0	74.3	58.3	1.5 (0.8–2.7)	$P = .19$ (NS; chi-square test)
Unemployed = 1	25.7	41.7		
Income (%)				
CDN $40,000 or more = 0	60.6	16.7	3.6 (1.4–9.2)	$P = .000^{\ddagger}$ (chi-square test)
CDN $39,000 or less = 1	39.4	83.3		
Age (mean and SD)	29.4 (9.0)	26.7 (9.0)		$P = .2$ (NS; Student t test)
Length of treatment (weeks) (mean and SD)	11.6 (6.4)	21.9 (20.3)		$P = .009^{\ddagger}$ (Student t test)
Pain duration (months) (mean and SD)	47.4 (53.8)	41.6 (45.7)		$P = .6$ (NS; Student t test)
No. of treatments (mean and SD)	1.8 (0.8)	2.3 (0.6)		$P = .02^{\dagger}$ (Student t test)
Treating clinician (% improvement)				
Clinician 1	50.0	50.0		
Clinician 2	61.9	38.1		$P = .9$ (NS; chi-square test)
Clinician 3	60.0	40.0		
Clinician 4	50.0	50.0		

*Critical RR = 2.0.
$^{\dagger}P < .05$; $^{\ddagger}P < .001$.
Tests used: chi-square test for differences between proportions; Student t test for differences between means, critical relative risk = 2.0.

10.0 ($P < .005$). In addition, the ability to group similar items into a semantic cluster was greater in the rTMD group (27 ± 11) than in the nrTMD group (18 ± 7.4) ($P < .001$). The CCC, which assessed immediate memory recall in the presence of a verbal interference, was also able to differentiate between the rTMD and nrTMD populations, with the former demonstrating greater recall ability than the latter (36 ± 6.3 versus 30 ± 6.3, respectively; $P < .01$).

Psychosocial Tests

With the SAQ, the nrTMD patients reported scores that were 20% higher than those of the rTMD patients, reflecting a sleep disturbance ($P < .05$). According to the BDI, there were no statistically significant differences between groups, despite the fact that there were substantially higher depression scores in patients with nrTMD (41% higher) as compared to those patients with rTMD. The lack of statistical significance may be attributed to the sample size required for this particular instrument. Fatigue levels were 48% higher for nrTMD ($P < .01$) as compared to rTMD patients. Alternatively, energy levels were 14% lower for the nrTMD patients, but this also was not a statistically significant difference.

Clinical Examination

Analysis of clinical parameters was also done to determine their utility in predicting treatment outcome. Only 1 of these parameters demonstrated a statistically significant predictive capacity: pain on palpation of the posterior ligament of the TMJ in the external auditory meatus (TMJEAM). In fact, 40% of the rTMD patients had positive scores for TMJEAM, while 74% of nrTMD patients had a positive score (RR: $P < .05$).

Assessment of the Role of Confounders for Logistic Regression Analysis

As shown in Table 2, 8 potential confounders that could not be eliminated during the design of the study were analyzed. Moreover, since the primary purpose of the study was to assess the presence of pretreatment predictors of patient outcome and not to determine the "best" modality of treatment, it was felt that "treatment" in and of itself could be considered a single "entity." As part of the

analysis, the success rates and number of modalities used, on average, by each treating clinician were assessed separately. As illustrated in Table 2, there was no difference in the number of modalities utilized or the treatment success rates for any clinician, thereby supporting the decision to handle "treatment" as a single entity even though different combinations and permutations of treatment modalities may have been used by any 1 clinician (ie, all clinicians were equally "successful" in obtaining positive outcomes). Moreover, the clinicians were blinded as to the baseline measures for all tests and could not alter treatment on the basis of poor or good scores on those tests. With respect to additional potential confounders, there was no difference in educational level or employment between the 2 TMD groups. However, the proportion of low-income individuals in the nrTMD group was significantly higher ($P < .0001$) than that in the rTMD group. Notably, the length of treatment for patients with nrTMD was almost twice that for patients with rTMD (21.9 weeks versus 11.6 weeks, respectively; $P < .01$). Analogously, the number of treatments provided was also greater for patients with nrTMD (2.3 versus 1.8; $P < .05$), which is consistent with earlier studies that focused on posttraumatic TMD.[6,18] The remaining variables (ie, age, pain duration, and treating clinician) were not different when the 2 populations were compared. Therefore, for the purposes of logistic regression analysis, the following confounders were included: *(1)* educational and income level, *(2)* length of treatment, and *(3)* number of treatments provided. As noted above, there was no difference in the mean age of each study population. However, age is traditionally considered a confounder in most epidemiologic studies, and therefore, despite our negative findings, this factor was also included in the logistic regression analyses.[30]

In addition to the confounders selected from Table 2, other psychosocial variables that have been reported to be associated with neuropsychologic tests (eg, fatigue, energy level, sleep disturbance) were included in regression analyses. We also included depression in the analysis, even though it did not reach a statistically significant level as a predictor in our study, because the depression level was marginally significant in the nrTMD group and also since this is thought to influence the other tests being used. It was thought that, if anything, the inclusion of depression in our analyses would bias against the demonstration of positive findings in relation to predictive tests. Finally, any variable that involved pain experience (eg, pain at rest pretreatment, pretreatment pain duration, and pretreatment pain on chewing) was included. Therefore, it was determined that none of the confounders included in the logistic regression analysis influenced the association between neuropsychologic test scores and treatment outcomes described on Table 1.

Impact of Neuropsychologic Tests as Predictors of Treatment Outcome

The data shown in Table 3 suggested that data from individual neuropsychologic and psychosocial tests predicted treatment outcome for TMD. These tests have published norms, which were indistinguishable from the internal control values shown in our study, demonstrating their high reproducibility.[16,23,25,26] To assess the overall impact of these parameters relative to their ability to predict treatment outcome, the following tests were dichotomized (ie, placed into "responder" and "nonresponder" categories): *(1)* CVLT-CR, *(2)* the CCC, *(3)* the SAQ, and *(4)* the BDI. Depression, because of its apparent relevance regarding treatment outcome for TMD, was also included.[9,10,12,25] Other tests, the data of which were positive predictors, such as the CVLT-CL and fatigue, were excluded because published norms were not available. Moreover, their high standard deviations made it impossible to define a reasonable cutoff point. From Table 3, it was determined that patients who had positive test results on the CCC and SAQ were 2.2 and 3.1 times more likely to become nonresponders, respectively. However, the findings were nonsignificant for the CVLT-CR (1.8 times for nonresponse) and BDI (1.7 times for nonresponse), probably because the sample size estimates used were based on detecting differences in means rather than proportions, which can be easily noticed in the wide 95% confidence interval for all tests.[30]

Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were also calculated for all tests. The calculations were based on 2 disease populations (rTMD versus nrTMD). Scores of 0.75 can be considered good for all measures. None of the tests reached a "good" score for all 4 measures,[11] and only the CCC had better results than the CVLT-CR, with moderate to good sensitivity and specificity (0.58 and 0.76, respectively) as well as PPV and NPV (0.64 and 0.72, respectively).

Backward stepwise analysis was used to design a model utilizing the best predictors for TMD treat-

Table 3 Confounders in Responding and Nonresponding TMD Patients

Independent variables (unit or category)	Responding TMD group (n = 36) (%)	Nonresponding TMD group (n = 24) (%)	Relative risk (95% CI)*	Significance	Sensitivity	Specificity	PPV	NPV
CVLT-CR					0.16	0.94	0.66	0.63
Scores 45–80 = 0	94.4	83.3	1.8	$P = .2$ (NS;				
Scores 0–44 = 1	5.6	16.7	(0.9–3.4)	Fisher exact)				
CCC					0.58	0.76	0.64	0.72
Scores 31–45 = 0	76.5	41.7	2.2	$P = .007^{\dagger}$				
Scores 0–30 = 1	23.5	58.3	(1.2–4.2)	(chi-square)				
SAQ					0.91	0.46	0.47	0.85
Scores 0–16 = 0	31.4	8.3	3.1	$P = .03$*				
Scores 17–68 = 1	68.6	91.7	(0.8–11.5)	(chi-square)				
BDI					0.29	0.88	0.63	0.53
Scores 16–63 = 0	88.2	70.8	1.7	$P = .17$ (NS;				
Scores 0–15 = 1	11.8	29.2	(0.9–3.1)	Fisher exact)				

*$P < .05$; $^{\dagger}P < .01$.
Tests used: chi-square test for differences between proportions; Fisher exact test 2-tail for differences between proportions, critical relative risk = 2.
PPV = positive predictive value; NPV = negative predictive value; CVLT-CR = California Verbal Learning Test immediate recall; CCC = Peterson-Peterson Consonant Trigram Test; SAQ = Sleep Assessment Questionnaire; BDI = Beck Depression Inventory.

ment outcome. Analysis indicated that the best predictors among all parameters were the CCC score and presence of fatigue. The overall agreement (82.0%), sensitivity (0.78), and specificity (0.85) were all elevated. The CCC scores proved to be a good unconfounded predictor when other bivariate or multivariate analyses were used.

Discussion

Population, Recruitment, Follow-up Rates, and Pain Improvement

The social and demographic description of the TMD population described here is very similar to that of other TMD treatment outcome studies.[9,10,13,31] The VAS-based improvement rates (60%) and global change in symptoms variable-based (63%) improvement rates correlated very well with one another, as previously reported, but overall improvement was lower than that described in the literature (70% to 98%). We suggest that this was probably the result of the very restrictive inclusion/exclusion criteria used, as well as the strict criteria for improvement.[5,9,13,32–34] Moreover, this patient population was derived from a tertiary referral center, the implication being that management of such a group tends to be more difficult as a rule. It must also be emphasized that since the sample size was somewhat low, it is also conceivable that the outcomes reported in this study would have approached those reported in the literature more closely had a larger sample been obtained. Yet it must also be recognized that the ability to obtain statistically significant findings, even with a somewhat modest sample, could give more weight to the strength of the findings reported herein.

Predictors of Treatment Outcome for TMD Patients

Neuropsychologic Tests. Previous studies comparing patients with mild closed-head injuries to control subjects showed that these 2 populations could be differentiated on the basis of performance on neuropsychologic tests.[15] Other analogous studies[17] showed that when similar tests were used, idiopathic TMD patients could be separated from posttraumatic TMD patients. By utilizing similar methodology, we showed here that idiopathic TMD patients might also be separated into responding and nonresponding TMD groups based on the results of certain neuropsychologic tests.

On average, the nrTMD population performed worse on the cognitive tests than the rTMD patients. However, no differences were found in the reaction time tests between the 2 groups. The reaction time test results in this study were very similar to those published in the literature for the idiopathic TMD and nonpain populations, confirming the reproducibility of these tests.[17,22,29] Since there were no differences in reaction times between the rTMD and nrTMD patients, the reaction time tests were not considered useful as predictors of treatment outcome for an idiopathic (ie, nontraumatic) TMD pain population.

Data obtained from some of the neuropsychologic tests employed in this study that evaluated attention, short-term memory, and short-term memory under interference were significantly different in the rTMD versus the nrTMD populations (Table 1). Nonresponding TMD patients remembered fewer words from a given "shopping list," were not as proficient in semantic clustering of that given list, and were less able to recall a group of 3 letters when challenged with a verbal interference as compared to the rTMD group. This factor may indicate a relative deficit in memory and/or concentration ability in the nonresponding population. Therefore, not only did the memory test scores appear to be good predictors of treatment outcome, but a strategy aimed at improving these functions could conceivably assist in the treatment of this particular subset of pain patients, as may be accomplished by administration of cognitive-behavioral therapy as described by others.[35–38]

Psychosocial Variables. When the SAQ was employed, we found that sleep scores in this population were comparable to those for patients with a primary diagnosis of sleep apnea, periodic leg movements, and snoring (mean = 26.0, SD = 8.6).[26] The proportion of individuals within the study population as a whole with sleep disturbances reached 78%, which may seem surprising. Thus it must be emphasized that, although the SAQ has been validated against polysomnography,[26] it would be imprudent to suggest here that unequivocal diagnoses of frank sleep disturbances were made. Assessments at baseline indicated that 68.6% of rTMD patients and 91.7% of nrTMD patients had a sleep disturbance as rated by their SAQ score, which as noted above, would seem quite high in relation to the incidence of sleep disturbances in other conditions. In any case, there appeared to be a relationship between sleep disturbances and nonresponse. However, as reported by others,[39] correction of sleep disturbances may not lead to improvements in TMD pain.

Unlike other studies, depression was not found to be a predictor of treatment outcome.[40,41] However, the absolute difference was relevant and the *P* value was marginally significant (*P* = .08), indicating that the limited sample size might have been the reason for this result. Moreover, 18.3% of the overall TMD population studied here was categorized as depressed, which is almost identical to that described by Gerschman and colleagues,[10] (18%) who used the Hamilton Depression Scale. This would be consistent with the notion that depression is still an important factor. In addition, the prevalence of depression in the total TMD population in this study was also substantially higher than the 6% found in other pain-free populations, again suggesting that depression must still be considered as an important factor affecting recovery, despite our negative findings.[42] However, other neuropsychologic and psychosocial variables or parameters were shown to be predictors of treatment outcome, even with this smaller sample, suggesting that these other factors may be more profound predictors than depression.

Clinical Examination Variables as Predictors of Treatment Outcome. The results obtained from clinical assessment in this study regarding TMJ and masticatory muscle tenderness were very similar to those published by Dworkin and colleagues.[40,41] In those studies, dysfunctional TMD (ie, nrTMD) patients were indistinguishable from functional TMD (ie, rTMD) patients with respect to muscle tenderness as well as unassisted vertical range of jaw motion.[38,39] In our investigation, only 5% of the overall TMD population had maximal jaw opening of less than 35 mm. The only clinical predictor of poor treatment outcome was tenderness in the posterior ligament of the TMJ (via palpation from the external auditory meatus). However, the relevance of this finding is unclear, because the external auditory meatus is not actually within the anatomic boundaries of the TMJ.[43] In addition, it has been shown that dysfunctional TMD patients usually present positive responses to palpation of so-called placebo sites, which may be a factor.[44] Occlusal factors such as overbite and overjet did not have any predictive value either. Despite these conclusions, it must be kept in mind that the reliability of measurements concerning mandibular range of motion, as well as tenderness to palpation of the TMJ and masticatory muscles, is moderate. Therefore, these results must be interpreted with caution.[17,28,45]

As shown by others,[46] there were no differences in pretreatment pain at rest in either group. However, pretreatment pain on chewing was significantly higher (60 mm) in the nrTMD patients as compared to rTMD patients (39 mm). The average reduction in pain on chewing over baseline scores (36%) was significantly higher in the rTMD patients as compared to the nrTMD patients (0%). This inability to cope with pain during function has also been reported previously.[19,31,40,47]

Putative Mechanisms

At this time, the mechanisms underlying the association between chewing pain and nonresponse as well as the other findings pertaining to the predic-

tive nature of certain cognitive tests shown in this study are unclear. However, a possible neuropsychologic or neurophysiologic explanation may be considered. It is possible that pain with function may serve as a constant "reminder" of tenderness in the masticatory system. In relation to this, it has been speculated that the anterior cingulate cortex possesses both pain and attention/cognition centers within the cerebral cortex.[48] These regions receive input from the trigeminothalamocortical pathway, which terminates in the somatosensory cortex, which has been associated with the sensory-discriminative aspects of pain.[48] Activation of this pathway could explain why the nrTMD patients also perceive ongoing muscular and TMJ pain as well as pain on chewing when compared to the rTMD group, despite a lack of objective differences in clinical measures. It is possible that painful but otherwise noninjurious inputs from the muscles and joints cannot be "ignored" by patients with reduced cognitive function if they have pretreatment pain on chewing. Similarly, ongoing input to the somatosensory system provided by pain on chewing could cause continual activation of the anterior cingulate cortex, or perhaps in nonresponsive patients, this structure is continually in an active mode. Conceivably, this could lead to ongoing "recognition" of pain on the part of the patient. Such notions are quite speculative and are not directly supported by our data but provide interesting possibilities for future consideration and study. However, this study does provide some support for the biopsychosocial model for chronic pain.[49–51]

Future studies in this area may consider follow-up clinical and neuropsychologic assessment to determine the levels of improvement (or lack thereof) in each category. Our findings, as well as those of others,[47,52] may suggest that recovery depends not on elimination of muscle pain or limited opening but rather the ability of a cognitively intact individual to, in effect, "learn to ignore" the painful muscles and mandibular limitation (by bypassing, deactivating, or otherwise not activating the anterior cingulate cortex).[48,53,54] It is also important to point out that although cognitive factors do appear to be useful for prediction of treatment outcome, it is also clear that tests for these factors cannot replace a thorough history and examination. Nonetheless, the apparent link between cognition and pain should lead to further studies of cognition, attention, and pain centers in the central nervous systems of idiopathic TMD patients, which may lead to further understanding of the underlying pathophysiology of TMD and perhaps other chronic pain conditions.

Acknowledgments

The authors wish to thank Ms Violeta Tapia, who assisted in the preparation of this manuscript.

References

1. Bell WE. Orofacial Pain: Classification, Diagnosis, Management, ed 4. Chicago: Year Book Medical Publishers, 1989.
2. Bonica JJ. Chapter 5. In: Bonica JJ, Liebeskind JC, Albe-Fessard DG (eds). Advances in Pain Research and Therapy. Vol 3: Proceedings of the Second World Congress on Pain, Montréal, Québec, 1978. New York: Raven Press, 1979.
3. Bonica JJ. Chapter 6. In: Bonica JJ, Liebeskind JC, Albe-Fessard DG (eds). Advances in Pain Research and Therapy. Vol 3: Proceedings of the Second World Congress on Pain, Montréal, Québec, 1978. New York: Raven Press, 1973.
4. Bonica JJ. Chapter 7. In: Bonica JJ, Liebeskind JC, Albe-Fessard DG (eds). Advances in Pain Research and Therapy. Vol 3: Proceedings of the Second World Congress on Pain, Montréal, Québec, 1978. New York: Raven Press, 1973.
5. Greene CS, Laskin DM. Long-term evaluation of treatment of myofascial pain dysfunction syndrome: A comparative analysis. J Am Dent Assoc 1983;107:235–238.
6. Kolbinson DA, Epstein JB, Senthilselvan A, Burgess JA. A comparison of TMD patients with or without prior motor vehicle accident involvement: Treatment and outcomes. J Orofac Pain 1997;11:337–345.
7. Vallon D, Nilner M, Soderfeldt B. Treatment outcome in patients with craniomandibular disorders of muscular origin: A 7-year follow-up. J Orofac Pain 1998;12:210–218.
8. Fricton JR, Kroening RF, Hathaway KW. Temporomandibular Joint and Craniofacial Pain; Diagnosis and Management. St Louis: Ishiyaku EuroAmerica, 1988.
9. Fricton JR, Olsen T. Predictors of outcome for treatment of temporomandibular disorders. J Orofac Pain 1996;10:54–65.
10. Gerschman JA, Wright JL, Hall WD, Reade PC, Burrows GD, Holwill BJ. Comparisons of psychological and social factors in patients with chronic orofacial pain and dental phobic disorders. Aust Dent J 1987;32:331–335.
11. Dworkin SF, LeResche L. Research Diagnostic Criteria for Temporomandibular Disorders. J Craniomandib Disord Facial Oral Pain 1992;6:301–355.
12. Gale EN, Funch DP. Factors associated with successful outcome form behavioral therapy for chronic temporomandibular joint pain. J Psychosom Res 1984;28: 441–448.
13. Lipton JA, Marbach JJ. Predictors of treatment outcome in patients with myofascial pain-dysfunction syndrome and organic temporomandibular joint disorders. J Prosthet Dent 1984;51:387–393.
14. Pankhurst CL. Controversies in the aetiology of temporomandibular disorders. Part 1. Temporomandibular disorders: All in the mind? Prim Dent Care 1997;4:25–30.
15. Stuss DT, Ely P, Hugenholtz H, et al. Subtle neuropsychological deficits in patients with good recovery after closed head injury. Neurosurgery 1985;17:41–47.

16. Stuss DT, Stethem LL, Poirier CA. Comparison of three tests of attention and rapid information processing across six age groups. Clin Neuropsychol 1987;1:139–152.
17. Goldberg MB, Mock D, Ichise M, et al. Neuropsychologic deficits and clinical features of posttraumatic temporomandibular disorders. J Orofacial Pain 1996;10:126–140.
18. Romanelli GG, Mock D, Tenenbaum HC. Characteristics and response to treatment of posttraumatic temporomandibular disorders: A retrospective study. Clin J Pain 1992;8:6–17.
19. Dao TT, Lavigne GJ, Charbonneau A, Feine JS, Lund JP. The efficacy of oral splints in the treatment of myofascial pain of the jaw muscles: A controlled clinical trial. Pain 1994;56:85–94.
20. Price DD, Bush FM, Long S, Hawkins SW. A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales. Pain 1994;56:217–226.
21. Folkard S, Glynn CJ, Lloyd JW. Diurnal variation and individual differences in the perception of intractable pain. J Psychosom Res 1976;20:289–301.
22. Stuss DT, Stethem LL, Hugenholtz H, Richard MT. Traumatic brain injury: A comparison of three clinical tests, and analysis of recovery. Clin Neuropsychol 1989;3:145–156.
23. Delis D, Kramer JH, Kaplan E, Ober BA. California Verbal Learning test. Psychological Corporation, 1987.
24. Peterson LR, Peterson MJ. Short-term retention of individual verbal items. J Exper Psychol 1959;58:193–226.
25. Beck AT. Depression: Causes and Treatment. Philadelphia: University of Pennsylvania Press, 1970.
26. Cesta A, Moldofsky H, Sammut C. The University of Toronto sleep assessment questionnaire (SAQ). Sleep Res 1996;25:486.
27. Taylor W. How to use guide to sample size calculation. In: Taylor W, (ed) Hamilton, Ontario, Canada: McMaster University Press, 1981.
28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.
29. Stuss DT, Stethem LL, Hugenholtz H, Picton T, Pivik J, Richard MT. Reaction time after head injury: Fatigue, divided and focused attention, and consistency of performance. J Neurol Neurosurg Psychiatry 1989;52:742–748.
30. Hennekens CH, Buring JE. Epidemiology in Medicine. Boston: Little, Brown, 1987.
31. Schnurr RF, Rollman GB, Brooke RI. Are there psychologic predictors of treatment outcome in temporomandibular joint pain and dysfunction? Oral Surg Oral Med Oral Pathol 1991;72:550–558.
32. Magnusson T, List T, Helkimo M. Self-assessment of pain and discomfort in patients with temporomandibular disorders: A comparison of five different scales with respect to their precision and sensitivity as well as their capacity to register memory of pain and discomfort. J Oral Rehabil 1995;22:549–556.
33. Millstein-Prentky S, Olson RE. Predictability of treatment outcome in patients with myofascial pain-dysfunction (MPD) syndrome. J Dent Res 1979;58:1341–1346.
34. Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. Pain 1983;16:87–101.
35. Whitney CW, Dworkin SF. Practical implications of noncompliance in randomized clinical trials for temporomandibular disorders. J Orofac Pain 1997;11:130–138.
36. Dworkin SF. Behavioral and educational modalities. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1997;83:128–133.
37. Oakley ME, McCreary CP, Clark GT, Holston S, Glover D, Kashima K. A cognitive-behavioral approach to temporomandibular dysfunction treatment failures: A controlled comparison. J Orofac Pain 1994;8:397–401.
38. Greco CM, Rudy TE, Turk DC, Herlich A, Zaki HH. Traumatic onset of temporomandibular disorders: Positive effects of a standardized conservative treatment program. Clin J Pain 1997;13:337–347.
39. DeNucci DJ, Sobiski C, Dionne RA. Triazolam improves sleep but fails to alter pain in TMD patients. J Orofac Pain 1998;12:116–123.
40. Dworkin SF, LeResche L, von Korff M, Truelove E, Sommers E. Predicting continued presence and level of temporomandibular disorder pain. J Dent Res 1989;68:194.
41. Dworkin SF, LeResche L, von Korff M, Dicker B, Sommers E. Constant, remitted and cyclic pain patterns in temporomandibular disorder: Three-year follow-up. J Dent Res 1991;70:441.
42. Nielsen AC, Williams TA. Depression in ambulatory medical patients: Prevalence by self-report questionnaire and recognition by non-psychiatric physicians. Arch Gen Psychiatry 1980;37:999–1004.
43. Laskin DM. Headache and the aching jaw [editorial]. J Oral Maxillofac Surg 1983;41:420.
44. Wilson L, Dworkin SF, Whitney C, LeResche L. Somatization and pain dispersion in chronic temporomandibular disorder pain. Pain 1994;57:55–61.
45. Carlsson GE, Egermark-Eriksson I, Magnusson T. Intra- and inter-observer variation in functional examination of the masticatory system. Swed Dent J 1980;4:187–194.
46. Linton SJ, Melin L. The accuracy of remembering chronic pain. Pain 1982;13:281–285.
47. Freeman BV, Psutka DJ, Hunter JJ, Baker GI. Arthroscopy of the temporomandibular joint: A comparison of objective and subjective outcome measures in patients with and without evidence of a psychopathologic disorder. Alpha Omegan 1998;91:44–50.
48. Davis KD, Taylor SJ, Crawley AP, Wood ML, Mikulis DJ. Functional MRI of pain- and attention-related activations in the human cingulate cortex. J Neurophysiol 1997;77:3370–3380.
49. Epker J, Gatchel RJ, Ellis E III. A model for predicting chronic TMD: Practical application in clinical settings. J Am Dent Assoc 1999;130:1470–1475.
50. Goldstein BH. Medical legal considerations in temporomandibular disorders. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 1999;88:395–399.
51. Talo S, Rytokoski U, Hamalainen A, Kallio V. The biopsychosocial disease consequence model in rehabilitation: Model development in the Finnish 'work hardening' programme for chronic pain [see comments]. Int J Rehabil Res 1996;19:93–109.
52. Tenenbaum HC, Freeman BV, Psutka DJ, Baker GI. Temporomandibular disorders: Disc displacements. J Orofac Pain 1999;13:285–290.
53. Sikes RW, Vogt BA. Nociceptive neurons in area 24 of rabbit cingulate cortex. J Neurophysiol 1992;68: 1720–1732.
54. Pillay PK, Hassenbusch SJ. Bilateral MRI-guided stereotactic cingulotomy for intractable pain. Stereotact Funct Neurosurg 1992;59:33–38.