Reliability, Validity, and Clinical Utility of the Research Diagnostic Criteria for Temporomandibular Disorders Axis II Scales: Depression, Non-Specific Physical Symptoms, and Graded Chronic Pain

Samuel F. Dworkin, DDS, PhD
Washington State Dental Foundation Distinguished Professor
Professor Emeritus
Departments of Oral Medicine and Psychiatry and Behavioral Sciences

Jeffrey Sherman, PhD
Acting Assistant Professor
Department of Oral Medicine

Lloyd Mancl, PhD
Research Assistant Professor
Department of Dental Public Health Sciences

University of Washington
Seattle, Washington

Richard Ohrbach, DDS, PhD
Associate Professor
Department of Oral Diagnostic Sciences
University at Buffalo
Buffalo, New York

Linda LeResche, ScD
Research Professor

Edmond Truelove, DDS, MSD
Professor and Chair

Department of Oral Medicine
University of Washington
Seattle, Washington

Correspondence to:
Dr Samuel F. Dworkin
University of Washington
Box 356370
Seattle, WA, 98195
Fax: 206-685-8412
E-mail: dworkin@u.washington.edu

Aims: *To analyze the reliability, validity, and clinical utility of the depression, non-specific physical symptoms, and graded chronic pain scales comprising the Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) Axis II.* ***Methods:*** *Data resulting from independent longitudinal and cross-sectional epidemiological studies as well as randomized clinical trials conducted at the University of Washington and the University at Buffalo were submitted to descriptive, correlational, and inferential statistical analyses to evaluate selected psychometric properties of the RDC/TMD Axis II scales.* ***Results:*** *Analyses of available data from both TMD clinical centers revealed good to excellent reliability, validity, and clinical utility for the Axis II measures of depression, somatization, and graded chronic pain. Specifically, data were presented comparing the RDC/TMD depression scale to the Beck Depression Inventory and the Center for Epidemiologic Studies Depression Scale; these data supported concurrent validity of the RDC/TMD measure and its use as a depression screening tool. Its clinical utility lies in its demonstrated usefulness for alerting TMD clinicians to potentially noteworthy depressive symptomatology in TMD patients. Others have shown that elevated somatization, the tendency to report non-specific physical symptoms as noxious or troublesome, is a predictor of poor TMD treatment outcome. The present analyses demonstrated that the RDC/TMD Axis II non-specific physical symptoms scale has acceptable reliability and that severe levels of somatization can potentially confound interpretation of the Axis I clinical examination. The graded chronic pain scale was demonstrated to have clinical utility for tailoring TMD treatment to levels of a patient's psychosocial adaptation.* ***Conclusion:*** *The major RDC/TMD Axis II measures demonstrate psychometric properties suitable for comprehensive assessment and management of TMD patients.*
J OROFAC PAIN 2002;16:207–220.

Key words: temporomandibular disorders, depression, somatization, chronic pain, RDC/TMD, reliability, validity

The Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) provide clinical researchers with a standardized system that can be evaluated for its use in examining, diagnosing, and classifying the most common subtypes of TMD.[1] Since their introduction in 1992, the RDC/TMD have been widely used in clinical research settings around the world where TMD and orofacial pain are managed.[2–4] Translations, in whole or in part, have been created and used in clinical studies with Dutch, Finnish, French, German, Hebrew, Japanese, Spanish,

and Swedish[5] populations; in addition, Chinese, Danish, Italian, Korean, Portuguese, and Romanian versions of the RDC/TMD have been produced but not yet field-tested. The original (ie, US) version of the RDC/TMD has also been recommended as a model system generalizable to investigating the diagnosis and classification of any chronic pain condition.[6]

The major attributes of the RDC/TMD that make them especially valuable in clinical research settings are: *(1)* a carefully documented and standardized set of specifications for conducting a systematic clinical examination for TMD; *(2)* operational definitions stated in unambiguous measurable terms for major clinical variables (eg, range of jaw motion, pain during muscle palpation, joint sounds); *(3)* demonstrated reliability for these operationally defined clinical measurement methods; and *(4)* use of a dual-axis system: Axis I to record clinical physical findings, and Axis II to record behavioral (eg, mandibular functional disability), psychologic (eg, depression, somatization), and psychosocial status (eg, chronic pain grade for assessing pain severity and life interference).[1]

Reliability of Measurement and Validity of Diagnosis

Until the present, with some important exceptions, the reliability, validity, and clinical utility of the Axis I and Axis II components of the RDC/TMD have not been thoroughly examined. While the diagnoses for other dental conditions are thought of as highly valid and reliable, demonstration of the validity of clinical tests for conditions such as TMD is extremely complex, because the condition does not yield many objective measurements. Most clinical findings from a TMD examination are the result of subjective reporting by the patient (eg, pain level during jaw motion, level of pain in response to muscle palpation) or subjective reporting by the clinician (eg, the subjective detection of joint sounds through digital examination).[7] This high degree of subjectivity has led us to suggest[1] the incorporation of multiple methods for increasing the reliability of TMD assessment and diagnosis. For example, we have advocated for repeated calibration of examiners to increase reliability. This is a strategy also advocated by the World Health Organization,[8] and Reit[9] found that even diagnoses of dental caries based on radiographic evidence can be unreliable when based on interpretations by noncalibrated examiners.

The relationship between reliability and validity is a critical one, both conceptually and psychometrically,[10] yet the obvious implication of this relationship is often overlooked in the TMD clinical literature. Conceptually, a valid diagnosis cannot result from the use of unreliable diagnostic examination methods, whether those methods involve use of physical instruments, clinical examination and testing procedures, or self-reported history and symptom data. The validity of the diagnosis—whether or not a person has a given condition—is limited by the reliability of the diagnostic methods used to obtain the clinical diagnosis. Because reliability in this case reflects consistency of findings from an examination procedure (ie, how well an examination procedure agrees with itself), there is a limit to how well the same examination procedure could agree with an external measure assessing validity. Thus, reliability can act as a limit to validity. Reliability of measurement is at the core of valid or useful diagnostic procedures, and if reliability is low, validity cannot be determined.

Reliability of RDC/TMD Axis I Measures

Axis I measures of the RDC/TMD are used to obtain a physical diagnosis of the most commonly occurring subtypes of TMD—specifically, masticatory muscle disorders, disc displacements, and inflammatory or degenerative disease of the temporomandibular joint (TMJ). Many clinical studies have demonstrated that without careful standardization and calibration, clinicians are not likely to achieve acceptable levels of reliability. Our own work has shown that assessment of clinical signs such as TMJ sounds can be highly unreliable without proper calibration of examiners.[11,12] As a result, the assessment of clinical findings typically is associated with inconsistency, ie, poor reliability.[13] In response, several methods have been developed for maximizing reliability of assessment for pain conditions. For example, Waddell and colleagues[14] have advocated the creation of clear and operationalized criteria for the assessment of clinical signs of back pain. We have argued for clear examination specifications and examiner calibration.[2,11] Prior studies have demonstrated that the Axis I component of the RDC/TMD is associated with acceptable to high levels of reliability when the clinical examination is performed as specified.[12,15–18]

Reliability of RDC/TMD Axis II Measures

Axis II measures of the RDC/TMD assess jaw disability during function, psychologic status, and psychosocial level of functioning. In contrast to the clinical examination measures of the type incorporated into Axis I, Axis II measures assess behavioral, psychologic, and psychosocial function through subjective self-report. Subjective self-report measures, such as published tests approved by the American Psychological Association (the Minnesota Multiphasic Personality Inventory,[19] the Beck Depression Inventory [BDI],[20] the Multidimensional Pain Inventory,[21] and the Symptom Checklist-90 [SCL-90]),[22] must meet published criteria for acceptable levels of reliability and validity.

The gold standard for psychiatric classification and diagnosis of mental and emotional disorders as established by the American Psychiatric Association is a highly structured interview, such as the "Diagnostic Interview Schedule" from the *Diagnostic and Statistical Manual of Mental Disorders*.[23] The use of such a formal psychiatric diagnostic measure has clearly demonstrated that clinic populations of TMD patients include many individuals who meet criteria for major depressive, anxiety, and somatization disorders.[24] However, such lengthy and structured assessments are impractical for the dental practitioner without training or time to administer such measures.

The RDC/TMD Axis II measures provide a reasonable surrogate for such lengthy diagnostic devices, because they were not intended to yield clinical psychiatric diagnoses. The ultimate validity question for the RDC/TMD Axis II is whether the relevant measures of psychologic status and level of psychosocial functioning are clinically useful in guiding TMD clinician actions and treatment planning rather than yielding formal psychiatric diagnoses for mental disorders.

Determinations of such Axis II components as depression, somatization, and chronic pain dysfunction are based on published psychologic tests or behavioral measures that have been determined to be reliable[15,21,24] and whose validity has been demonstrated by comparison to established measures. Such methods as carefully structured and lengthy psychiatric interviews and use of psychologic tests designed to diagnose major psychiatric disorders are clearly not useful in clinical dental settings, because dentists have neither the time nor the training to comfortably make formal psychiatric diagnoses.

Use of the RDC/Axis II Measures

The team of TMD specialists who contributed to the creation of the RDC/TMD Axis II[1] sought to include a set of measures for arriving at an integrated assessment of: *(1)* the physical status of the masticatory system, and *(2)* the extent to which TMD sufferers report psychologic distress, including depressive symptomatology and the presence of widespread non-specific physical symptoms, and psychosocial disturbance, such as the heavy toll TMD pain might be taking on their daily lives. This integration into the RDC/TMD of behavioral, psychologic, and psychosocial self-report measures was undertaken for TMD assessment to be consistent with prevailing biopsychosocial models for understanding disease and illness—approaches that use interdisciplinary methods in the management of all chronic pain conditions.[25] Assessing these latter (Axis II) factors is strongly advocated and routinely accomplished in all multidisciplinary pain centers because such assessment results in more rational treatment decisions and more realistic treatment expectations on the part of both the patient and the health care provider.

The specific measures in the psychologic and psychosocial domains incorporated into Axis II were chosen to serve as a standardized screen for significant emotional upset and TMD-related disruption of activities of daily living. The RDC/TMD Axis II measures are not intended to yield clinical psychiatric diagnoses. Instead, they assess the extent to which a person with TMD may be so cognitively, emotionally, or behaviorally impaired that these biobehavioral factors may contribute to the development or maintenance of the pain problem[26] and/or interfere with smooth acceptance of and compliance with treatment. In rarer instances, they alert the clinician to perceived threats to overall well-being and even to life through patient-initiated responses to straightforward questions.[27,28]

Assessment of the validity of Axis II measures includes examining how those measures of psychologic and psychosocial function compare to similar measures whose psychometric properties have been better established. For example, if an Axis II measure places the patient in the high range of severe psychologic symptoms, would comparable findings be obtained on another, well-studied psychologic test designed for the same purpose? Because the stated purpose of the Axis II measures is to classify patients into normal, moderate, or severe ranges of functioning on symptoms and behaviors indicating psychologic disturbance, one

aspect of the validity of such Axis II measures—concurrent validity—rests on the degree to which they compare with measures designed to accomplish the same purpose and that are administered at the same time. Another aspect of the validity of Axis II regards clinical utility, that is, the extent to which the Axis II measures can guide clinical decision-making.

The present report reviews issues underlying the reliability and validity of clinical tests, then focuses on assessment of the reliability, validity, and clinical utility of RDC/TMD Axis II measures of depression and non-specific physical symptoms and the Graded Chronic Pain Scale. For each Axis II measure, data are presented concerning *(1)* reliability, as assessed by Cronbach's alpha, a well-known measure of the internal consistency of a psychologic scale; *(2)* concurrent validity, that is, agreement between an Axis II measure and an independent measure of the same characteristic; *(3)* construct validity, or how well the test measures a theoretical construct; and *(4)* clinical utility, the effectiveness of using the Axis II instrument in a clinically meaningful way.

Materials and Methods

Subjects

The analyses reported are based on data derived from participants in 6 independent studies; these include a longitudinal epidemiologic study that identified TMD pain cases and TMD pain-free controls,[29,30] as well as several longitudinal clinical studies that identified consecutive cases appearing for treatment of TMD-related pain at different TMD and orofacial pain clinics located at the University of Washington (UW)[31–34] and the University at Buffalo (UB). Data from these several independent sources of TMD patients and data gathered on different samples at separate points in time are included in several of the analyses presented to examine more rigorously the reliability, validity, and clinical utility of these RDC/TMD Axis II measures.

The origin, composition, and size of each of the 6 samples analyzed in this study are as follows.

UW Databases:

- Population-based samples: *(1)* Community TMD cases (n = 113); *(2)* Pain-free community controls (n = 210)
- Clinic samples of TMD patients from tertiary care settings: *(3)* A database of consecutive patients with TMD appearing for treatment at the Orofacial Pain and Temporomandibular Disorders Clinic (n = 242); *(4)* Baseline data from clinic patients with TMD who participated in a randomized clinical trial of a dental hygienist–delivered treatment program (n = 124); *(5)* Baseline data from clinic patients who participated in a randomized clinical trial incorporating a psychologic treatment program into usual TMD care (n = 117)

UB Database:

- *(6)* A consecutive series of TMD cases appearing for treatment at the TMD and Orofacial Pain Clinic (n = 226)

Measures

RDC/TMD Axis II. The RDC/TMD Axis II measures to be reported included depression and somatization scales adapted from the SCL-90. The SCL-90 is a straightforward checklist of symptoms, and the patient simply indicates the extent to which they have been bothered by specific symptoms in the past month on a 0 to 4 scale. The depression and somatization scales have been widely used with chronic pain patients.[2] Normative data defining cutoff scores for normal, moderate, and severe levels of depression and numbers of non-specific physical symptoms, commonly referred to by the label "somatization," were provided by a large population-based study.[1]

Comparison Measures of Depression. Additional measures of depression, not included in the RDC/TMD Axis II, were included in the present study to allow initial assessment of the concurrent validity, sensitivity, and specificity of the RDC/TMD Axis II measure of depression. These comparison measures included the following.

Beck Depression Inventory (BDI).[20] The BDI consists of 21 groups of items that assess both cognitive/affective and neurovegetative symptoms of depression. It is widely used to measure depression symptoms in chronic pain. A cutoff score of 16 is commonly used to separate those with low versus high levels of depressive symptomatology.[15]

Center for Epidemiologic Studies Depression Scale (CES-D).[35] The CES-D is a 20-item self-report inventory designed to assess depressive symptomatology. Respondents are asked to indicate how frequently they experienced each symptom in the past week, ranging from 0 (less than 1 day) to 3 (5 to 7 days). The total possible score ranges from 0 to 60 and reflects both the number of symptoms and the frequency of their occur-

rence. The internal consistency of the CES-D has been reported to be between 0.84 to 0.90.[35]

Psychosocial Function. The RDC/TMD Axis II measures the level of psychosocial function through the use of the Graded Chronic Pain Scale (GCP),[36] a self-report instrument designed to provide a quantitative index for assessing the impact and severity of chronic pain. The GCP provides a meaningful quantitative index integrating perceived pain intensity and the extent to which pain is psychosocially disabling. Disability is measured by extent of pain-related interference with daily activities and number of lost activity days (eg, days unable to go to work or school, attend to household responsibilities) attributed to TMD pain. Grade I is defined as TMD pain of low intensity, averaging less than 5 on a 10-point scale, and associated with little pain-related interference in daily living. Grade II is defined as high-intensity pain (above 5 on a 10-point scale) with low amounts of pain-related interference. Grades III and IV are associated with increasing levels of pain-related psychosocial disability regardless of pain level. For most analyses previously reported, we have defined grades I and II patients as not disabled by their TMD pain whether pain is severe or mild, while grades III and IV patients, typically combined in our analyses because of their low prevalence, reflect moderate to significant pain-related psychosocial disability, which is almost always accompanied by severe pain intensity.

Procedures

All subjects in the population-based studies completed a questionnaire that included the RDC/TMD Axis II measures; in addition, all identified TMD cases underwent an RDC/TMD Axis I clinical examination. All TMD patients also completed the questionnaires, described above, included in the respective clinical studies in which they participated. All data reported were baseline RDC/TMD Axis I and Axis II data gathered according to published protocols for the use of the RDC/TMD and before any clinical treatment was initiated. For selected analyses, some post-treatment data are also presented. Reliability data were calculated by the use of Cronbach's alpha[37] for all estimates and concurrent validity by a calculation of Pearson correlations between the RDC/TMD measure and comparable measures administered at the same time. Where feasible, construct validity and predictive validity were also assessed; for example, construct validity of the nonspecific physical factors scale was determined by the use of an exploratory factor analysis with a principal-components factor extraction method, and the predictive validity of Varimax rotation was assessed by a comparison of outcomes of randomized clinical trials based on baseline GCP assessment. Clinical utility was assessed by an examination of RDC/TMD Axis II findings (eg, numbers of nonspecific physical symptoms) and their relationship to clinical findings (number of muscle sites tender to palpation on an RDC/TMD Axis I physical examination).

Results

Depression

Reliability. The reliability of the RDC/TMD depression scale was examined to assess its internal consistency, as summarized in Table 1. Data used were from population-based studies and from a randomized clinical trial that gathered data from clinic cases over several years. The Cronbach's alpha values for internal consistency of the RDC/TMD Axis II scale of depression ranged from 0.91 to 0.93, indicating excellent reliability, regardless of whether the assessment was part of an epidemiologic study, from a sample of clinic cases, or from several independent samples of TMD cases from population-based or clinical research studies gathered at widely separated points in time (see Table 1).

Concurrent Validity. Concurrent validity coefficients were calculated to estimate the correlation between scores on the RDC/TMD Axis II measure for depression and those on the BDI and the CES-D. Data for these analyses came from 2 separate samples of TMD cases for which RDC/TMD Axis II data were available. One sample constituted baseline data from a cohort of UW clinic cases for whom BDI data were also available (n = 115); a second data set was available from clinic cases for whom CES-D data were also available (n = 186). Concurrent validity was $r = 0.69$ for the RDC/TMD Axis II depression scores when BDI scores were the standard for comparison and r = 0.78 when CES-D scores were the standard; both values indicate good levels for confirming concurrent validity.

Clinical Utility. The assessment of the clinical utility of diagnostic measures includes assessment of sensitivity—the ability of the diagnostic measure to detect a disease condition when it is truly present—and specificity—the ability of the measure to not diagnose a disease as present when there is no disease. The underlying principle

Table 1 Reliability (Cronbach's alpha) of RDC/TMD Axis II Measures of Depression, Non-Specific Physical Symptoms, and Graded Chronic Pain

	Chronbach's alpha levels		
RDC/TMD Axis II measure	UW clinic and community cases (n = 362)	UW TMD clinic cases (RCTs) (n = 242)	UB TMD clinic cases (n = 226)
Depression	0.91	0.93	0.92
Non-specific physical symptoms			
Total scale	0.82	0.86	0.87
Pain items excluded	0.78	0.83	0.80
Graded chronic pain	0.71	0.90	N/A

RCTs = randomized clinical trials.

Table 2 Clinical Utility of the RDC/TMD Axis II Measure of Depression

	RDC Axis II depression score (% of cases)		
	Normal	Moderate	Severe
UW TMD clinic cases *(n = 115; P < .001)*			
versus BDI normal (BDI ≤ 16) (n = 69)	40.6	39.1	20.3
versus BDI depressed (BDI > 16) (n = 46)	8.7	13.0	78.3
UB TMD clinic cases *(n = 186; P < .001)*			
versus CES-D normal (CES-D ≤ 16) (n = 59)	69.5	16.9	13.6
versus CES-D high (CES-D > 16) (n = 127)	18.9	29.9	51.2

assumed when assessing sensitivity and specificity of a diagnostic measure is that some standard of truth, a so-called gold standard, is available by which we can know that a disease or other pathologic state is truly present. Because the purpose of these analyses was to establish clinical utility and not to validate the depression scale of RDC/TMD Axis II for its ability to accurately diagnosis clinical depressive disorders, the comparisons of the RDC/TMD Axis II depression measure with the BDI and the CES-D were extended to include an examination by chi-square analysis of the extent of association between them as well as determination of sensitivity and specificity for the RDC/TMD depression measures when either the BDI or the CES-D was selected as the gold standard for comparison purposes.

Table 2 presents chi-square analyses of the distribution of "normal," "moderate," and "severe" depression scores derived from the RDC/TMD Axis II measure of depression, compared to TMD cases above and below conventionally accepted cutoff criteria for depression associated with the BDI and the CES-D. For example, when the recommended BDI scores above and below cutoffs for depression on the BDI were compared to the RDC/TMD Axis II categories, of those subjects classified as depressed by the BDI measure, 78% also scored as "severe" by the RDC/TMD measure; 22% were misclassified by the RDC/TMD if the BDI was taken as the standard for actual depression. Converted to sensitivity and specificity, these data indicate that when the BDI is the gold standard, the Axis II measure of depression, collapsed into 2 categories of "normal" versus "moderate/severe" depression scores, has a sensitivity of 0.91 and specificity of 0.41. Compared to the CES-D data, which were derived from a different sample of TMD cases, sensitivity is 0.81 and specificity is 0.70, ie, both measures showed greater sensitivity than specificity (ie, were more likely to detect patients who might not be truly depressed than to miss patients who are depressed). This reinforces our contention that the RDC/TMD Axis II depression measure is useful as a screen for depression but not as a diagnostic instrument for major depressive disorders.

Table 3 RDC/TMD Axis II Non-Specific Physical Symptoms Scale Factor Structure (Principal Components) for TMD Cases and Controls

	Full scale, TMD cases 2-factor solution (n = 324)		Pain items excluded, TMD cases 1-factor solution (n = 324)	Full scale, TMD pain-free controls 5-factor solution (n = 204)				
Non-specific physical symptom scale items	**Factor 1:** Pain and fatigue (35.4% of variance)	**Factor 2:** Cardio-pulmonary (10.5% of variance)	**Factor 1:** Non-specific physical (36.0% of variance)	**Factor 1** (23.3% of variance)	**Factor 2** (12.29% of variance)	**Factor 3** (9.9% of variance)	**Factor 4** (9.3% of variance)	**Factor 5** (8.7% of variance)
Pain and fatigue								
Headaches	**0.727**	–0.239		0.148	0.113	0.174	–0.027	**0.769**
Faintness	**0.580**	0.150	0.557	0.231	0.399	0.272	–0.135	**–0.527**
Pain in lower back	**0.575**	0.170		0.293	0.177	**0.548**	–0.320	0.191
Nausea/upset stomach	**0.584**	0.191		0.089	0.279	**0.744**	–0.055	0.111
Sore muscles	**0.539**	0.326		**0.825**	–0.144	0.259	0.014	0.021
Weakness in body parts	**0.570**	0.479	**0.783**	**0.822**	0.174	–0.053	–0.047	0.205
Heavy feeling in arms/legs	**0.532**	0.492	**0.748**	**0.554**	0.138	0.200	0.148	–0.300
Cardiopulmonary								
Pain in chest	0.016	**0.748**		0.069	0.102	0.022	**0.844**	0.014
Trouble getting breath	0.123	**0.809**	**0.660**	0.231	0.399	0.272	–0.135	**–0.527**
Non-specific symptoms								
Hot or cold spells	0.482	0.382	**0.630**	0.029	**0.713**	0.033	0.337	0.063
Numbness or tingling	0.486	0.411	**0.648**	0.124	–0.177	**0.701**	0.227	–0.151
Lump in throat	0.297	0.496	**0.586**	0.181	**0.719**	–0.082	–0.294	–0.180

Items with factor loadings > 0.50 indicated in bold type.

Non-Specific Physical Symptoms

The RDC uses scale items derived from the widely used somatization scale of the SCL-90 to assess the tendency to report non-specific physical symptoms as noxious or troublesome. Five items assess the presence of different kinds of pains (Table 3). For some analyses it seems reasonable to exclude these pain items, making it clear that TMD patients may experience a wide variety of physical symptoms, apart from the extent to which they report widely dispersed pain symptoms.

Reliability. Table 1 includes data on reliability, assessed by Cronbach's alpha, for the Axis II non-specific physical symptom scale, with and without pain items included. The differences in alpha levels shown for the non-specific physical symptom scale, with and without pain items, reflect in part the influence of scale length on determining reliability; all other things being equal, a shorter scale will be less reliable than a scale containing a larger number of comparable items. Good to acceptable levels of internal consistency were observed for both scales with and without pain items.

Concurrent/Construct Validity. Because there are no readily available well-documented scales to assess the tendency to report non-specific physical symptoms (ie, reliable and valid somatization scales), it was not possible to relate Axis II non-specific physical symptom scale scores to a concurrent alternative measure with known psychometric properties, as we were able to do when we compared the Axis II depression scale scores with the BDI or the CES-D (see above). Instead, we examined the construct validity of the Axis II scale of non-specific physical symptoms by conducting a factor analysis of that scale on TMD cases and non-TMD controls. The resulting principal components analysis, as summarized in Table 3, yielded a 2-factor solution when the entire scale (including pain items) was used for TMD cases with factor loadings > 0.50 applied as the criterion for including an item in a particular factor. (see Table 3, bold type factor loadings)

- Pain and Fatigue: This factor accounted for 35.4% of the total variance, and included all the pain items in the scale except pain in the chest,

Table 4 Mean No. of Painful Muscle Palpation Sites for UW TMD Clinic and Community Cases (n = 342) Scoring Normal, Moderate, and Severe on RDC/TMD Axis II Non-Specific Physical Symptoms Scale

Symptom scale score	n	Mean no. of sites	SE	95% confidence intervals	F	*P*
Normal	130	5.7	0.5	4.7–6.9	27.25	.0001
Moderate	103	8.3	0.8	7.0–9.5		
Severe	91	12.1	0.8	10.6–13.6		

plus faintness and heavy feelings in the arms and legs.

- Cardiopulmonary: This second factor, accounting for 10.5% of total variance, included only 2 items, relating to cardiac pain and breathing difficulties.

Repeating the factor analysis with the same TMD cases but with the pain items excluded from the scale, a single factor emerged, as seen in Table 3, encompassing all but 5 of the scale items and accounting for 36% of variance.

By contrast, when the factor analysis was repeated on non-TMD pain-free controls, no useful factor structure emerged. As Table 3 shows, the 12 scale items are dispersed among more than 5 factors with no apparent pattern. These data support the possibility that non-specific physical symptom reporting arises systematically in samples of TMD cases, but is not a systematic issue with TMD pain-free controls. Results from extensive factor analyses of the non-specific physical symptom scale of the RDC/TMD, based on independent clinical samples at UB,[38] essentially cross-validated the factor analytic results summarized in Table 3.

Clinical Utility. It would be clinically useful to know if the tendency to self-report many non-specific physical symptoms, as reflected in RDC/TMD Axis II scale scores, might interact with those portions of the RDC/TMD Axis I clinical examination that require self-reporting of pain or other symptoms. For example, is the tendency to self-report widespread, unrelated physical symptoms related to heightened self-report of pain as numerous masticatory muscles are palpated during the course of an Axis I diagnostic clinical examination? The mean number (and 95% confidence limits) of muscle sites painful to palpation during the Axis I clinical examination is shown in Table 4, based on data from the UW epidemiologic studies involving clinic and community TMD cases. The potential range for the number of extraoral muscle sites tender to palpation is 0 to 20. The mean number of painful muscle sites (with virtually no overlap in distributions) is 5.7 for those in the normal range versus 12.1 ($P < .0001$) for those in the "severe" range on the RDC/TMD Axis II non-specific physical symptoms scale.

For the UW series of TMD cases, Table 4 summarizes scores of normal, moderate, or severe levels of non-specific physical symptoms and number of muscles painful to palpation on RDC/TMD standardized clinical examination. From these data, it was possible to determine, for example, that only 20% of TMD patients at the normal level on the Axis II non-specific physical symptom scale had more than 10 painful palpation sites, while of those scoring severe, 2.6 times more, or about 52%, self-reported 10 or more painful masticatory muscle sites. Only 1 severe-scoring TMD patient, or 1% of those scoring severe, reported no painful Axis I palpation sites, compared to 21% of those scoring normal.

Further confirming the clinical utility of the RDC/TMD somatization scale is the previously reported observation,[39] drawn from the same epidemiologic data, of a strong association between low (< 0.50), moderate (0.50 to 0.99), or high (> 1.00) scores on the somatization scale and number of defined placebo sites reported as painful during an RDC/TMD Axis I standardized clinical examination. Forty-five percent of those scoring high reported 1 or more placebo sites on the face and head painful to palpation, compared to only 15% of those scoring low on somatization (chi-squared, $P < .005$).[39]

Table 5 summarizes the relationship between somatization scale score and number of painful muscles that were examined for the UB sample of TMD clinic cases, with the use of RDC/TMD cut-

Table 5 Mean No. of Painful Muscle Palpation Sites for UB Clinic Cases (n = 128) Scoring Normal, Moderate, and Severe on RDC/TMD Axis II Non-Specific Physical Symptoms Scale

Symptom scale score	n	Mean no. of sites	SE	95% confidence intervals	F	*P*
Normal	47	3.4	0.6	2.2–4.6	11.78	.0001
Moderate	34	5.9	0.8	4.4–7.4		
Severe	47	8.6	1.0	6.7–10.5		

off points for normal, moderate, and severe somatization scores. The results were similar to those observed with the UW sample. For example, from the UB data shown in Table 5, it was possible to determine that while 65% (ie, nearly two-thirds) of those scoring normal on the non-specific physical symptom scale had 3 or fewer painful masticatory muscle sites, only 19% of those scoring severe reported 3 or fewer painful masticatory muscle sites. At the other end of the clinical continuum, from the UB dataset it was possible to determine that more than 7 painful masticatory muscle sites were reported by 62% of those scoring severe on the non-specific physical symptom scale, compared to only 24% of those scoring in the normal range.

Graded Chronic Pain

The RDC/TMD Axis II classification level of psychosocial function is captured by the GCP, developed from our longitudinal TMD studies.[36]

Reliability. Reliability has been reported previously and is incorporated into data from the original publication of the GCP,[36] which demonstrate that the GCP has acceptable to good reliability of 0.71. Based on data sets from more recent randomized clinical trials, Cronbach's alpha averaged 0.90; this indicates a high internal consistency (see Table 1).

Concurrent Validity. Concurrent validity assessment again utilized previously reported data from large-scale epidemiologic studies, which confirm the validity of the GCP as a hierarchical scale, where each level of GCP score (from 0 to IV) is associated with escalating TMD pain-related disability. Pain-related disability may be expressed as impact on activities of daily living, unemployment, utilization of health care and medications, depression, and poor self-perceived health status. These psychosocial parameters of chronic pain disability were used as criteria to validate the properties of the GCP by determining that increases in pain grade were associated with significant increases in numbers of patients reporting disability (Fig 1). For example, with regard to the impact of TMD pain on daily life, only 5.1% of grade I TMD patients reported high impact; this increased to 25.4%, 63.4%, and 71.4% for grades II, III, and IV, respectively.

Clinical Utility. The clinical utility of the GCP was tested (see Table 6) by conducting 2 randomized clinical trials, which tailored treatments for TMD patients according to GCP status, relatively independent of Axis I physical diagnosis.[32,33] Two very different treatment protocols were used: the first was for those with a GCP score of grade I or grade II-Low (indicating no disability points, ie, persons not heavily impacted psychosocially by their TMD problem), and another was for those of grade II-High (indicating pain-related interference) or grades III or IV (indicating elevated pain-related interference). Each treatment protocol compared the tailored treatment intervention to the usual care delivered to TMD patients in the Orofacial Pain Clinic, UW Department of Oral Medicine, by attending dentists/TMD specialists.

In 1 randomized clinical trial, those scoring I and II-Low (ie, minimal pain-related disability/interference) were randomly assigned to a manual-based TMD self-management treatment protocol delivered by a trained registered dental hygienist over 3 sessions plus 2 telephone contacts, in lieu of the usual TMD care provided by a TMD specialist.[40] The self-management protocol was designed as a programmatic approach by which these TMD patients could learn the concepts and skills needed to become adept at self-monitoring and self-care for their TMD condition. These patients, we hypothesized, could do just as well with less than the usual amount of TMD treat-

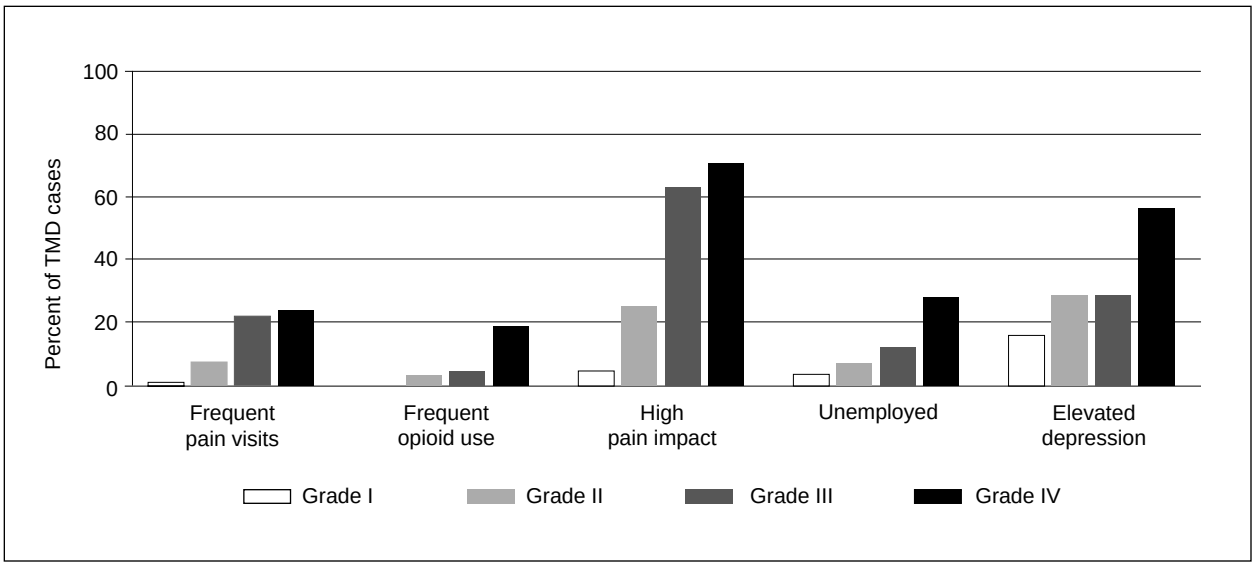

Fig 1 Concurrent validity for RDC/TMD Axis II graded chronic pain scale (GCP): Current level of psychosocial disability by GCP (I to IV). Data taken from TMD clinic cases (n = 347).

ment. In a separate randomized clinical trial, those with a GCP score identified as not adapting well to their TMD—that is, grades II-High, III, and IV—were randomly assigned to a protocol that supplemented the usual treatment by a clinic dentist/TMD specialist with a 6-session cognitive-behavioral intervention. The intervention introduced many of the same self-care notions that were incorporated into the hygienist-based clinical trial, but it also included an intense component aimed at modifying pain-related cognitions, emotions, and behaviors. These poorly adapting pain patients, we hypothesized, needed more than the usual treatment, but did not require more biomedically based treatment.

Table 6 shows that in each randomized clinical trial, where the intervention of a hygienist in lieu of the usual treatment or a psychologist to complement the usual treatment was compared with the respective usual treatment as a control, those patients in the intervention arm tailored to their GCP level did better at the post-treatment evaluation than those who received usual care. The major dependent variables in Table 6 included pain intensity, as measured by visual analog scales, and self-reported ability to control pain.

Discussion

Data have been presented that support the reliability, validity, and clinical utility of measurement and patient classification in 3 domains assessed by the RDC/TMD Axis II: *(1)* depression, *(2)* presence of non-specific physical symptoms, and *(3)* GCP. Our conclusions are summarized in Table 7.

We have shown that Axis II measures of depression, somatization, and GCP have acceptable to excellent reliability, validity, and clinical utility. With regard to depression, data were presented comparing the RDC/TMD depression scale to the BDI and the CES-D. Our data support the concurrent validity of the RDC/TMD measure and its use as a screening measure of depression. Its use is not yet supported as a diagnostic instrument because it has not been tested for its ability to provide a valid diagnosis of a clinical depressive disorder according to criteria elicited in a structured clinical interview. Its clinical utility lies in its demonstrated usefulness for alerting clinicians to the potential existence of noteworthy depressive symptomatology in their TMD patients.

Somatization, the tendency to report non-specific physical symptoms as noxious or trouble-

Table 6 Clinical Utility of RDC/TMD Axis II Graded Chronic Pain Scale (GCP) for TMD Treatment Decisions

	Post-treatment means					
	RCT 1			RCT 2		
Post-treatment outcome measure	Self management (RDH) (n = 61)	Usual treatment (DDS) (n = 63)	*P*	Comprehensive care (CBT + DDS) (n = 59)	Usual treatment (DDS) (n = 58)	*P*
Average pain intensity (0–10 VAS)	2.9	3.1	.53	4.4	5.6	.02
Ability to control pain (0–5 VAS)	3.8	3.2	.07	4.1	3.1	.001

Summary of 2 randomized controlled trials (RCTs). RCT 1: GCP I, II-Low were assigned to self-management program in lieu of usual treatment. RCT 2: GCP II-High, III, IV were assigned to comprehensive care (psychologist-led cognitive-behavioral therapy integrated with usual treatment).

RDH = registered dental hygienist; DDS = dentist; CBT = cognitive-behavioral therapy.

Table 7 Summary of Results of Analyses Conducted to Assess Reliability, Validity, and Clinical Utility of RDC/TMD Axis II

Axis II measure	Reliability	Concurrent validity	Clinical utility
Depression	Excellent	Good to acceptable	• Good sensitivity and low specificity • Utility comparable to BDI and CES-D, as screening instrument only, for depressive symptomatology and depressive disorder
Non-specific physical symptoms	Good	(Construct validity): Comparisons of factor structures in clinic and community and TMD pain-free controls	Clinical utility demonstrated for: • Screening potential over-reporting of positive responses to RDC/TMD Axis I clinical examination for painful masticatory muscle sites • Screening for report of pain as a non–TMD-specific physical symptom
Graded chronic pain	Good	Excellent	Utility demonstrated for contributing to evidence-based decision-making regarding tailoring of TMD treatment to individual level of psychosocial adaptation to TMD pain

some, has already been shown by others to be a predictor of TMD treatment outcome.[41] The RDC/TMD Axis II non-specific physical symptoms scale score was shown in the present report to have good reliability, for both the entire scale and the scale after removal of pain items. We have provided additional data indicating that severe levels of somatization have the potential for confounding interpretation of the Axis I clinical exam. We speculate that the association between reporting many non-specific physical symptoms and reporting many muscles painful to palpation may reflect, for some patients, heightened non-specific physiologic reactivity for detecting physical sensations in the body.[42–44] It would be helpful if clinical researchers, as well as clinicians, carefully exam-

ined the possibility that some subjective reports in response to the TMD physical clinical examination might represent an interaction between specific altered muscle or joint pathophysiology and/or heightened vigilance[45] for self-reporting many physical symptoms that are of a non-specific nature.

Finally, we have presented data on the clinical utility of the RDC/TMD Axis II measure of GCP to serve as a clinical decision-making tool for tailoring treatment regimens to the patient's level of psychosocial adaptation. Specifically, it would appear that TMD patients with low GCP scores, ie, grades I (low in pain and disability) and II-Low (high pain with low disability) might benefit from exposure to self-management programs emphasizing self-care and self-monitoring of TMD symptoms. For most of these patients, recurrence of symptoms in this clinical condition, which has been shown to have a recurrent, fluctuating course, should not indicate the need for escalation to more extensive, invasive, or expensive treatments, but a return to carefully structured self-management strategies proven to have been effective in earlier episodes.

There are, of course, many unaddressed issues associated with the data presented. First, the reliability, validity, and clinical utility data would need to be replicated on data sets representing multiple and diverse sets of patients. While some of the data analyzed in the present report came from different sources and were gathered at different points in time, a more systematic examination of populations of TMD patients, including gathering of longer follow-up data, is warranted to confirm conclusions presented earlier. In addition, Axis II of the RDC/TMD contains measures of jaw function and anxiety, both of which require much more examination of their respective reliability and validity than has been conducted to date.

It is also noted that test-retest reliability has not been performed sufficiently enough to warrant reporting at present, largely due to patient and personnel constraints. While interexaminer reliability seems a much more critical concern, it is nevertheless very useful to provide test-retest reliability estimates to determine consistency of individual examiners over time. Such reliability analyses are already planned for implementation as part of a large-scale U.S. National Institutes of Health–funded multicenter in-depth study of all critical aspects of the reliability and validity of the RDC/TMD.

Because TMD still lack clear and useful gold standards to assure the diagnosing clinician that, for example, a true chronic muscle pathology exists or not, we are in fact diagnosing by description. In this context it is also useful to remember that, although our present TMD classification systems are not based on underlying pathologic states observed through objective measurement, our patients nevertheless experience very real pain, meaningful discomfort, limitations in jaw function, and, too often, elevated pain-related psychologic upset and psychosocial disturbance of their everyday lives.

It may also be helpful to be reminded once again that descriptive classification systems are how the most commonly occurring forms of other pain conditions, including tension headache,[46,47] back pain,[48] and fibromyalgia,[49] are classified. All of these conditions are plagued, in our current state of knowledge, by the absence of virtually any scientifically measurable physiologic parameters that could, singly or in combination, be used to verify a pathophysiologic state. This state of affairs does not mean there are no reliable and valid pathophysiologic markers that represent the disease; rather, it means that we have not yet discovered what those gold standard criteria are (eg, valid radiographic criteria for TMJ pathology linked to pain, or biologic assays or magnetic resonance imaging findings validating muscle pain as arising from pathophysiology of masticatory muscles).

Our data support the use of the RDC/TMD Axis II components as reliable and valid indicators of depression, somatization, and psychosocial dysfunction in response to pain. They do not support the use of the Axis II to provide psychiatric diagnoses. Rather, they provide initial scientific support to validate evidence-based clinical decision-making. If further clinical research provides additional evidence of the clinical utility of incorporating findings from Axis II into treatment plans, this may lead to more comprehensive approaches to the treatment of TMD. For example, if the depression scale of Axis II is demonstrated to be clinically useful—valid for the clinician to use as described in the RDC/TMD—then a response to the patient's report of severe depressive symptomatology may guide the dentist to an expansion of treatment responsibilities. The latter may include office-based TMD treatment methods that take into account the depressive symptoms.

There is increasing awareness in health care that depression may best be first detected and management instituted in primary care, as opposed to specialty or tertiary health care settings.[50] Alternatively, use of and attention to the Axis II measure of depression may result in recognition

and referral to a mental health care provider by the dental practitioner. Such treatment options would be desirable if it could be shown that elevated depressive symptomatology in TMD patients is associated with excessive use of health care, undue reliance on medications, and reporting of multiple non-specific physical symptoms spread throughout the body, all of which represent a scientifically established risk factor for poor TMD treatment outcome. It is in this sense that we approach validating the Axis II of the RDC/TMD as having clinical utility, because its component measures demonstrate, with known levels of statistical certainty, the absence or presence of psychologic disturbance and/or psychosocial dysfunction and because, in turn, the patient's status within these Axis II psychosocial parameters can have direct consequences for the course of their TMD condition and its management.

Acknowledgments

This research was supported by NIDCR, grant number DE-10766.

References

1. Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders: Review, criteria, examinations and specifications, critique. J Craniomandib Disord Facial Oral Pain 1992;6:301–355.
2. Goulet J-P, Lavigne GJ, Lund JP. Jaw pain prevalence among French-speaking Canadians in Quebec and related symptoms of temporomandibular disorders. J Dent Res 1995;74:1738–1744.
3. List T, Dworkin SF. Comparing TMD diagnoses and clinical findings at Swedish and U.S. TMD centers using research diagnostic criteria for temporomandibular disorders. J Orofac Pain 1996;10:240–253.
4. Lobbezoo-Scholte AM, DeLeeuw JRJ, Steenks MH, Bosman F, Buchner R, Olthoff LW. Diagnostic subgroups of craniomandibular disorders part I: Self-report data and clinical findings. J Orofac Pain 1995;9:24–36.
5. Dworkin SF. Surveying the feasibility of an RDC/TMD-based International TMD Research Consortium [abstract]. J Dent Res 2001;80:55.
6. Garofalo JP, Wesley AL. Research Diagnostic Criteria for Temporomandibular Disorders: Reflection of the physical-psychological interface. APS Bull 1997;May/June:4–16.
7. Ohrbach R, Dworkin SF. Five-year outcomes in TMD: Relationship of changes in pain to changes in physical and psychological variables. Pain 1998;74:315–326.
8. World Health Organization. Oral Health Surveys—Basic Methods. Geneva: World Health Organization, 1971.
9. Reit C. On decision making in endodontics: A study of diagnosis and management of periapical lesions in endodontically treated teeth. Swed Dent J 1986;Suppl 41:1–30.
10. Cronbach LJ, Glaser GC, Handa H, Rajaratna S. The Dependability of Behavioral Measurements. New York: Wiley, 1972:
11. Dworkin SF, LeResche L, DeRouen T. Reliability of clinical measurement in temporomandibular disorders. Clin J Pain 1988;4:89–99.
12. Dworkin SF, LeResche L, DeRouen T, Von Korff M. Assessing clinical signs of temporomandibular disorders: Reliability of clinical examiners. J Prosthet Dent 1990;63:574–579.
13. Fleiss JL, Chilton NW. The measurement of interexaminer agreement on periodontal disease. J Periodont Res 1983;18:601–606.
14. Waddell G, Main C, Morris EW, et al. Normality and reliability in clinical assessment of backache. Br Med J 1982;284:1519–1523.
15. Dworkin SF, Sherman JJ. Relying on objective and subjective measures of chronic pain: Guidelines for use and interpretation. In: Turk DC, Melzack R (eds). Handbook of Pain Assessment, ed 3. New York: Guilford Press, 2001:619–638.
16. Rudy TE, Turk DC, Brody MC. Quantification of biomedical findings in chronic pain: Problems and solutions. In: Turk DC, Melzack R (eds). Handbook of Pain Assessment, ed 2. New York: Guilford Press, 1992: 447–472.
17. Rudy TE. Psychophysiological assessment in chronic orofacial pain. Anesth Prog 1990;37:82–87.
18. Clark JP, Flack VF. Reproducibility of examiner performance for muscle and joint palpation in the temporomandibular system following training and calibration. Community Dent Oral Epidemiol 1993;21:72–77.
19. Hathaway SR, McKinley JC. Manual for Administration and Scoring of the MMPI. Minneapolis: National Computer Systems, 1983.
20. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry 1961;4:561–571.
21. Turk DC, Rudy TE. The robustness of an empirically derived taxonomy of chronic pain patients. Pain 1990;43:27–35.
22. Derogatis LR, Cleary PA. Confirmation of the dimensional structure of the SCL-90: A study in construct validation. J Clin Psychol 1997;33(4):981–989.
23. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, ed 4. Washington, DC: American Psychiatric Association, 1994.
24. Gatchel RJ, Garofalo JP, Ellis E, Holt C. Major psychological disorders in acute and chronic TMD: An initial examination. J Am Dent Assoc 1996;127:1365–1374.
25. Bonica JJ. Basic principles in managing chronic pain. Arch Surg 1977;112:783–788.
26. Von Korff M, LeResche L, Dworkin SF. First onset of common pain symptoms: A prospective study of depression as a risk factor. Pain 1993;53:251–258.
27. Dworkin SF, Massoth DL. Temporomandibular disorders and chronic pain: Disease or illness? J Prosthet Dent 1994;72:29–38.
28. Turk DC. Strategies for classifying chronic orofacial pain patients. Anesth Prog 1990;37:155–160.
29. Dworkin SF, Huggins KH, LeResche L, et al. Epidemiology of signs and symptoms in temporomandibular disorders: Clinical signs in cases and controls. J Am Dent Assoc 1990;120:273–281.

30. Von Korff M, Dworkin SF, LeResche L, Kruger A. An epidemiologic comparison of pain complaints. Pain 1988;32:173–183.
31. Dworkin SF, Turner JA, Wilson L, et al. Brief group cognitive-behavioral intervention for temporomandibular disorders. Pain 1994;59:175–187.
32. Dworkin SF, Turner JA, Mancl L, et al. A randomized clinical trial of a tailored comprehensive care treatment program for temporomandibular disorders. J Orofac Pain (in press).
33. Dworkin SF, Huggins KH, Wilson L, et al. A randomized clinical trial using research diagnostic criteria for temporomandibular disorders-axis II to target clinic cases for a tailored self-care TMD treatment program. J Orofac Pain 2002:16(1):48–63.
34. Huggins KH, Truelove EL, Dworkin SF, Mancl L, Sommers E, LeResche L. RCT of splints for TMD: Clinical findings at 12 months [abstract 1491]. J Dent Res 1999;78(special issue):292
35. Radloff L. The CES-D scale: A self-report depression scale for research in the general population. J Appl Psychol Meas 1977;1:3385–3401.
36. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain 1992;50:133–149.
37. Bandura A. Self efficacy: Toward a unifying theory of behavior change. Psychol Rev 1977;84:191–215.
38. Ohrbach R, Dworkin SF, Truelove E. Domains of measurement in chronic TMD pain: Psychometric properties [abstract]. J Dent Res 1999;78:292.
39. Wilson L, Dworkin SF, Whitney C, LeResche L. Somatization and pain dispersion in chronic temporomandibular pain. Pain 1994;57:55–61.
40. Dworkin SF, Huggins KH, Wilson L, et al. A randomized clinical trial using research diagnostic criteria for temporomandibular disorders-axis II to target clinic cases for a tailored self-care treatment program. J Orofac Pain 2002;16(1):48–63.
41. McCreary CP, Clark GT, Oakley ME, Flack V. Predicting response to treatment for temporomandibular disorders. J Craniomandib Disord Facial Oral Pain 1992;6:161–169.
42. Rollman GB, Lautenbacher S. Sex differences in musculoskeletal pain. Clin J Pain 2001;17:20–24.
43. Jones D, Rollman GB, Brooke RI. The cortisol response to psychological stress in temporomandibular dysfunction. Pain 1997;72:171–182.
44. Maixner W, Fillingim R, Kincaid S, Sigurdsson A, Harris B. Relationship between pain sensitivity and resting arterial blood pressure in patients with painful temporomandibular disorders. Psychosom Med 1997;59:503–511.
45. McDermid AJ, Rollman GB, McCain GA. Generalized hypervigilance in fibromyalgia: Evidence of perceptual amplification. Pain 1996;66:133–144.
46. International Headache Society Classification Committee. Classification and diagnostic criteria for headache disorders, cranial neuralgias, and facial pain. Cephalalgia 1988;8(Suppl 7):1–96.
47. Solomon S. Diagnosis of primary headache disorders. Validation of the International Headache Society criteria in clinical practice. Neurol Clin 1997;15(1):15–26.
48. International Association for the Study of Pain Subcommittee on Taxonomy, Merskey H. Classification of chronic pain: Descriptions of chronic pain syndromes and definitions of pain terms. Pain 1986;(Suppl 3):S1–S225.
49. Wolfe F, Smythe HA, Yunus MB, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. Arthritis Rheum 1990;33(2):160–172.
50. Von Korff M. Pain management in primary care: An individualized stepped-care approach. In: Gatchel RJ, Turk DC (eds). Psychosocial Factors in Pain: Evolution and Revolution. New York: Guilford Publications, 1999:360–373.