Validating the TMJ Scale in a National Sample of 10,000 Patients: Demographic and Epidemiologic Characteristics

Stephen R. Levitt, MD, PhD
Clinical Associate Professor of Psychiatry
University of North Carolina at Chapel Hill
Director of Research
Pain Resource Center, Inc
Durham, North Carolina

Michael W. McKinney, PhD
Vice President
Pain Resource Center, Inc

Correspondence to:
Stephen R. Levitt, MD, PhD
Director of Research
Pain Resource Center, Inc
PO Box 2836
Durham, North Carolina 27715

The accuracy and reliability of the TMJ Scale were originally determined in cross-validation studies on large, research-based patient samples. It had been assumed that the demographic characteristics and test responses of these research-based samples would be representative of the clinical population in which the TMJ Scale would ultimately find use. The present study on more than 10,000 patients that were evaluated for temporomandibular disorders in clinical practice demonstrates that the test scores, demographic variables, and the patterns of symptom severity that characterize the original TMJ Scale research sample accurately represent the general temporomandibular disorder patient population in which the TMJ Scale is now being used. The results suggest a high degree of confidence in the clinical efficacy of this assessment tool. The overall symptom severity of temporomandibular disorders was found to be normally distributed in the patient population. Women with temporomandibular disorders report a higher level of severity of all physical and psychological symptoms than men. This may explain the high female-to-male ratio in patients seeking treatment. However, a higher percentage of male temporomandibular disorder patients has clinically significant psychological and stress-related problems than do women. The severity and prevalence of symptoms associated with joint dysfunction and range of motion limitation are lower in older age groups, and the overall symptom severity of temporomandibular disorders is not higher in older age groups. However, the severity and prevalence of symptoms associated with joint dysfunction are greater in groups in which temporomandibular disorders have existed for longer durations, although pain levels do not follow this trend. There is also an association between time duration of the temporomandibular disorder and the severity of psychological problems and chronicity. Patients with chronic problems are symptomatically more impaired than those with acute problems.
J OROFACIAL PAIN 1994;8:25–35

In 1982, the President's Council on Temporomandibular Disorders of the American Dental Association called for the development of screening questionnaires for temporomandibular disorders (TMD).[1] Partially in response to that request, a large collaborative research program was organized with the goal of developing a diagnostic aid for TMDs based on the symptom report of the patient.[2] The research program began in 1982 and spanned 5 years. A main objective was to develop a symptom-based diagnostic tool that could detect both the presence (sensitivity) and absence (specificity) of TMDs with sufficient accuracy for

the tool to be employed in actual clinical practice. A further goal was for the tool to be a truly multidimensional assessment aid by including a set of subscales to measure the severity of both the target physical symptoms of TMDs as well as psychological factors and stress.[2]

Models for the development of such assessment tools were well established during the past 50 years in the fields of psychology and behavioral science.[2-4] The most successful such models were chosen and adapted for the specific needs of a TMD assessment tool.[2] In 1986, after 5 years of research,[2,5-9] a tool named the TMJ Scale (Pain Resource Center) was introduced into clinical practice. Since that time, approximately 40,000 patients being evaluated and treated for TMDs have been tested with the TMJ Scale. Subsequent studies of the test's predictive value and clinical applications have demonstrated it to be clinically and statistically robust.[10-18]

The TMJ Scale responses are analyzed by a computer program.[2] Approximately 25% of dentists and other clinicians using this tool choose to send the patient's completed answer booklet to the test publisher's mail-in scoring service. At the time of this writing, more than 10,000 such tests have been scored by the test publisher. The availability of this large patient sample presents a unique opportunity for epidemiologic research.

Prior to release for general clinical use, large cross-validation clinical trials were performed to study the sensitivity and specificity of the TMJ Scale, as well as various types of statistical reliability.[2,6,7,9] These studies have demonstrated sensitivity and specificity in the 80% to 90% range for detecting TMDs in dental populations, as well as high levels of internal reliability (coefficient α) and test-retest reliability. Those studies were based on a large heterogeneous sample of 1,215 dental patients located in various regions throughout the US and Canada, of which 742 had been diagnosed as having a TMD.[2,9] A nonpatient sample of 1,191 normal individuals was used for comparison.[2,9] The demographic characteristics of the TMD population sample were assumed to be representative of the larger, clinical TMD population in which the test would ultimately find its use.[2,9] A unique opportunity now exists to study the demographics of a very large national sample of patients tested in actual clinical practice in every state in the US, as well as Canada, to determine if, in fact, this assumption was correct. It is now also possible to compare the distribution of test scores and structural validity of the tool derived from the original cross-validation research sample with the new national patient sample. The availability of a large national patient sample also allows some interesting epidemiologic issues to be explored in the TMD population. One can examine differences in severity of symptoms between men and women, in different age groups, and for the varying lengths of time during which a TMD has been present.

Materials and Methods

The first 10,549 completed TMJ Scale tests on patients 13 years of age and older consecutively received by the test publisher for scoring were accumulated in a research file for data acquisition and study. This patient sample is large and heterogeneous, collected from all 50 states and Canada, with 97.6% of the patients coming from dentists and the remainder from physicians, psychologists, and other health professionals. All demographic and test scoring information was taken directly from these completed tests and used in this study. Calculations were performed using the Statistical Package for the Social Sciences.[19]

Results

Two key demographic variables are the gender and age distributions in the two samples. Table 1 compares the gender and age characteristics in the cross-validation research sample with that found in the national clinical sample. The gender distributions in the two samples are remarkably similar, with the cross-validation research sample composed of 85.5% women and the national clinical sample of 85.1%. The age distributions in the two samples also manifest a very close match, with an average age of 33.6 years in the cross-validation research sample and 34.1 years in the national clinical sample. A median age of 32 years and range of 13 to 96 years in the cross-validation research sample are closely matched by a median age of 33 years and age range of 13 to 92 years in the national clinical sample. The actual breakdown into age subcategories in the two population samples shows that these two samples are very well matched (Table 1).

Three other demographic variables, marital status, education level, and ethnic-racial group, were examined (Table 1). The cross-validation research sample showed 29.9% of the patients to be single versus 31.6% in the national clinical sample. There was little variation between the two samples in each of the other marital status categories. Similar results were found for education level, with the average

number of years of schooling at 13.6 years in the cross-validation research sample and 13.5 years in the national clinical sample. The cross-validation research population was 94.4% white compared to 90.7% in the national clinical sample. Blacks made up 4.4% of the cross-validation research sample and 4.3% of the national clinical sample. Hispanics represent 0.7% of the cross-validation research sample and 3.0% of the national clinical sample.

Table 1 also shows the distribution of TMD duration in the two patient samples. The two groups were very similar. If one chooses 6 months as a cut-off to define a chronic disorder, then 22.1% of the cross-validation sample had chronic TMDs compared to 26.05% of the national sample.

Table 2 shows the comparison of average scores in the two population samples on each of the 10 subscales of the TMJ Scale. A close similarity in magnitude between the two patient samples is evident. The percentage differences between the average scores in the two groups, using the national clinical sample as reference, are shown in Table 2. For example, the Global Scale score mean was 1.80 in the cross-validation research sample and 1.83 in the national clinical sample. This represents a difference of 1.6%.

Table 3 shows the Pearson correlation matrix for the original cross-validation research sample and for the national clinical sample. Both may be compared to the correlation matrix of Spearman correlation coefficients for clinician ratings on the same symptom scales in a group of 803 patients.[2]

The most notable finding in Table 3 is the close correspondence between the magnitude of all coefficients. For example, the correlation coefficient between joint dysfunction and range of motion limitation is 0.50 in the cross-validation research sample compared with 0.57 in the national clinical sample. Using clinicians' ratings as the gold standard for such interrelationships, a Spearman correlation coefficient of 0.48 is found.[2] The absolute magnitudes of the various correlation coefficients in the two patient samples were very close, as were the various trends among the subscales. For example, in determining how joint dysfunction correlates with other physical symptoms of TMDs, joint dysfunction correlates highest with range of motion limitation, next highest with pain report, next highest with perceived malocclusion, and least with palpation pain (Table 3). This holds true in both the original cross-validation research sample and in the national clinical sample.

The TMJ Scale allows measurement and study of the distribution of TMD symptom severity within the TMD patient population. Figure 1 shows the frequency distribution curve for the numbers of patients from the national clinical sample achieving various symptom severity scores on the Global Scale.

Table 1 Demographics of the Cross-Validation Research Sample and the National Clinical Sample

Demographic Variable	Cross-Validation Research Sample (N = 742)	National Clinical Sample (N = 10,549)
Sex (%)		
Men	14.5	14.9
Women	85.5	85.1
Age (y)		
Mean	33.6	34.1
SD	13.2	13.4
Median	32	33
Range	13–96	13–92
Age subcategories (%)		
13–19 years	12.1	10.4
20–29 years	31.6	28.8
30–39 years	29.2	32.2
40–49 years	14.8	17.3
50–59 years	6.8	7.1
60+ years	5.5	4.2
Marital status (%)		
Single	29.9	31.6
Married	57.2	54.9
Separated	2.0	2.0
Divorced	6.6	7.6
Widowed	2.2	1.5
Remarried	2.1	2.4
Education		
Mean (y)	13.6	13.5
SD	2.8	2.7
Ethnic-racial group (%)		
Black	4.4	4.3
Hispanic	0.7	3.0
Asian	0.4	0.8
White	94.4	90.7
Other	0.1	1.2
Duration of TMD (%)		
<1 month	7.8	5.7
1–5 months	14.3	20.3
6–11 months	13.9	14.0
1–2 years	20.6	20.6
3–5 years	21.4	17.7
6–10 years	10.5	9.6
>10 years	11.5	12.9

One interesting epidemiologic issue is whether there is a difference between men and women in the perceived severity of TMD symptoms. Table 4 shows the average scores on the various subscales of the TMJ Scale for both men (N = 1,579) and women (N = 8,970) in the national clinical sample, along with the results of a *t* test of the means in the two groups.

Table 2 Comparison of Scores Between the Cross-Validation Research Sample and the National Clinical Sample in the Various Symptom Categories of the TMJ Scale

TMJ Scale symptom category	Cross-Validation Research Sample (N = 742)	National Clinical Sample (N = 10,549)	
	Mean (SD)	Mean (SD)	% Difference
Pain report	1.69 (0.89)	1.89 (0.93)	10.6
Palpation pain	1.13 (0.92)	1.39 (0.97)	18.7
Perceived malocclusion	1.69 (1.00)	1.70 (1.02)	0.6
Joint dysfunction	1.46 (0.92)	1.52 (0.93)	3.9
Range of motion limitation	2.09 (0.98)	2.15 (1.01)	2.8
Non-TMD	0.98 (0.68)	1.12 (0.70)	12.5
Psychological factors	1.02 (0.78)	1.15 (0.82)	11.3
Stress	1.37 (0.85)	1.58 (0.86)	13.3
Chronicity	0.80 (0.76)	0.90 (0.79)	11.1
Global scale	1.80 (0.75)	1.83 (0.78)	1.6

Table 3 Pearson Correlation Matrix for the Symptom Categories of the TMJ Scale as Determined in the Cross-Validation Research Sample (N = 742) and the National Clinical Sample (N = 10,549)

Scale	PR	PP	MO	JD	RL	NT	PF	ST	CN
PR	–	.77 (.80)	.38 (.44)	.26 (.42)	.58 (.68)	.71 (.72)	.41 (.35)	.43 (.40)	.49 (.48)
PP		–	.30 (.35)	.15 (.30)	.38 (.51)	.65 (.65)	.36 (.32)	.35 (.33)	.48 (.47)
MO			–	.18 (.31)	.41 (.51)	.37 (.39)	.27 (.23)	.35 (.31)	.30 (.26)
JD				–	.50 (.57)	.16 (.30)	–.02* (.10)	.05* (.16)	.04* (.10)
RL					–	.32 (.43)	.11 (.16)	.16 (.23)	.19 (.23)
NT						–	.53 (.51)	.53 (.50)	.59 (.63)
PF							–	.69 (.68)	.53 (.56)
ST								–	.42 (.42)
CN									–

( ) = National Clinical Sample. All correlations are significant to $P < .001$ except *not significant. PR = Pain Report; PP = Palpation Pain; MO = Perceived Malocclusion; JD = Joint Dysfunction; RL = Range of Motion Limitation; NT = Non-TMD; PF = Psychological Factors; ST = Stress; CN = Chronicity.

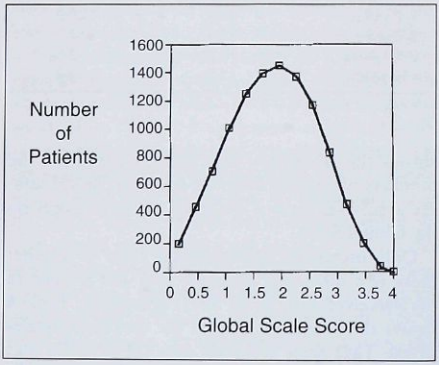

Fig 1 Frequency distribution of TMD symptom severity as measured by scores on the Global Scale of the TMJ Scale in a national sample of 10,549 patients.

Analyzing average scores on symptom scales is useful for investigating differences or trends between groups but can sometimes hide certain other important information. This approach does not indicate the numbers of individuals whose scores have reached clinical significance. To accomplish this, patients' scores may be compared to previously determined cut-off scores indicating clinical significance.[2] The sensitivity and specificity associated with these cut-off scores have been calculated for each subscale of the TMJ Scale,[2,9] and the results in the national clinical sample are shown in Table 5.

Another epidemiologic issue concerns whether the patient's age is associated with variations in severity or prevalence of symptoms. In this study, 4,139 patients were below the age of 30 years, 3,390 patients were between the ages of 30 and 39

Table 4 Comparison of Test Scores and *t* Test Between the Means for Men and Women on the TMJ Scale in a National Clinical Sample (N = 10,549)

Scale	Men (N = 1,579) Mean (SD)	Women (N = 8,970) Mean (SD)	*t* test *P* value
Pain report	1.65 (.96)	1.93 (.91)	.000
Palpation pain	1.19 (.94)	1.42 (.97)	.000
Perceived malocclusion	1.52 (1.00)	1.73 (1.01)	.000
Joint dysfunction	1.34 (.92)	1.55 (.93)	.000
Range of motion limitation	1.84 (1.02)	2.20 (1.00)	.000
Non-TMD	1.01 (.72)	1.14 (.70)	.000
Psychological factors	1.10 (.87)	1.15 (.81)	.034
Stress	1.40 (.86)	1.61 (.86)	.000
Chronicity	.88 (.81)	.90 (.79)	.317
Global scale	1.59 (.80)	1.87 (.77)	.000

Table 5 Percent of Patients With Clinically Significant Symptom Severities on the TMJ Scale According to Gender, Age, and Duration of TMD in a National Clinical Sample (N = 10,549)

	Gender		Age			Time Duration		
Scale	Males	Females	<30 (y)	30–39 (y)	>40 (y)	1–5 (m)	6 (m)–2 (y)	>3 (y)
PR	80.6	86.7	85.6	86.6	85.0	84.2	86.2	86.7
PP	70.7	89.4	86.7	87.1	85.7	83.9	97.8	87.7
MO	66.8	66.0	64.1	68.0	66.7	61.7	65.4	70.1
JD	67.4	75.4	79.5	73.4	67.8	70.1	74.7	77.0
RL	56.3	56.5	59.7	56.3	52.2	55.8	58.4	55.5
NT	63.4	62.2	60.2	64.7	62.7	58.0	62.2	66.1
PF	56.5	45.0	42.9	49.0	49.3	43.8	46.3	49.4
ST	71.4	54.7	53.0	61.3	58.3	50.4	55.0	64.7
CN	71.3	74.8	68.5	76.5	79.7	71.2	75.2	76.0

PR = Pain Report; PP = Palpation Pain; MO = Perceived Malocclusion; JD = Joint Dysfunction; RL = Range of Motion Limitation; NT = Non-TMD; PF = Psychological Factors; ST = Stress; CN = Chronicity.

years, and 3,020 patients were aged 40 years or older. Table 6 shows the average scores for these three age groups within the national clinical sample, along with the results of a *t* test of the means between various combinations of groups. Table 5 shows the percentage of patients in each age group with clinically significant levels of severity in the various symptom areas.

The national clinical sample was divided into three groups based on time duration of TMDs: 1 to 5 months (N = 3,074), 6 months to 2 years (N = 3,451), and 3 or more years (N = 3, 941). The severity of symptoms in these three groups is shown in Table 7. The percentage of patients in each group with clinically significant symptoms is shown in Table 5.

Discussion

The sensitivity, specificity, predictive values, and determination of reliability of the various subscales that make up the TMJ Scale have been previously published and were determined based on a large heterogeneous patient sample collected from across the US and Canada.[2,6,7,9,11–13]

A specific goal of the original research design was to study the psychometric properties of the test in a patient sample that would be truly representative of the TMD clinical population in which the test would ultimately be used. It is therefore important to collect a large sample of this clinical population and carefully compare its demographic characteristics with the original cross-validation research sample.

The results demonstrate that in the two key demographic variables of gender and age, the original cross-validation research sample was truly representative of the general TMD clinical population in which the TMJ Scale would ultimately be employed (Table 1).

The results also demonstrate that the marital status, education, and ethnic-racial demographics of the original cross-validation research sample

Table 6 Comparison of Test Scores and *t* Test Between the Means on the TMJ Scale for Different Age Ranges in a National Clinical Sample (N = 10,549)

	A (N = 4,139) <30 (y)	B (N = 3,390) 30–39 (y)	C (N = 3,020) 40+ (y)	A vs B	B vs C	A vs C
Scale	mean	mean	mean	*P* value	*P* value	*P* value
Pain report	1.87	1.94	1.86	.001	.000	.503
Palpation pain	1.34	1.42	1.42	.001	.987	.001
Perceived malocclusion	1.64	1.74	1.74	.000	.952	.000
Joint dysfunction	1.67	1.50	1.33	.000	.000	.000
Range of motion limitation	2.24	2.14	2.04	.000	.000	.000
Non-TMD	1.08	1.16	1.13	.000	.051	.005
Psychological factors	1.05	1.19	1.22	.000	.139	.000
Stress	1.50	1.66	1.59	.000	.003	.000
Chronicity	.75	.95	1.04	.000	.000	.000
Global scale	1.84	1.85	1.78	.490	.000	.003

Table 7 Comparison of Test Scores and *t* Test Between the Means on the TMJ Scale for Different Time Durations of TMD in a National Clinical Sample (N = 10,549)

	A (N = 3,074) 1–5 (m)	B (N = 3,451) 6 (m)–2 (y)	C (N = 3,941) >3 (y)	A vs B	B vs C	A vs C
Scale	mean	mean	mean	*P* value	*P* value	*P* value
Pain report	1.83	1.93	1.90	.000	.278	.001
Palpation pain	1.34	1.42	1.40	.001	.314	.011
Perceived malocclusion	1.59	1.69	1.79	.000	.000	.000
Joint dysfunction	1.41	1.53	1.60	.000	.000	.000
Range of motion limitation	2.09	2.20	2.14	.000	.013	.033
Non-TMD	1.06	1.13	1.17	.000	.011	.000
Psychological factors	1.08	1.14	1.20	.002	.004	.000
Stress	1.39	1.55	1.75	.000	.000	.000
Chronicity	.78	.90	.98	.000	.000	.000
Global scale	1.75	1.86	1.86	.000	.954	.000

were accurate in representing the general TMD patient population in which the TMJ Scale is now being used. The clinical sample included more nonwhites, especially Hispanics, than the research sample. However, the percentages were relatively small. Further study of the test results in minority groups is indicated.

Another variable that characterizes a TMD patient population is the time that the disorders have been present. The differences between patients with acute as opposed to chronic problems are of interest to both researchers and clinicians, especially when pain is an important component of the disorder. The results show that the original cross-validation research sample was representative of the general TMD patient population in terms of time duration of TMDs (Table 1).

Overall, the demographic studies clearly demonstrate the validity of the assumption that the original cross-validation research sample represented the general TMD patient population in which the TMJ Scale would be used.

Another method of investigating whether the original cross-validation research sample represents the general clinical population is to compare the average scores in the two populations on each of the 10 subscales of the TMJ Scale. The mean scores in the two populations do not show much variation, as seen by the calculated percentage differences (Table 2). The mean scores in the national clinical sample consistently exceed to some extent those in the original cross-validation research sample. This suggests that patients in the national clinical sample have a general trend of greater severity

in all measured symptom areas relative to the original cross-validation research sample. Overall, the results suggest that the original cross-validation research sample was a good approximation of the general TMD clinical population in which the TMJ Scale would subsequently be used.

Another form of inquiry into the suitability of the original cross-validation research sample and into the structural integrity of the test itself involves examination of the matrices of Pearson correlation coefficients among the subscales of the TMJ Scale.[2] This analysis helps to determine whether the underlying structure of symptom patterns is consistent in the two groups; essentially, whether patients tend to have similar patterns of dysfunction. A study of the variations in correlation coefficients among all subscales reveals a close correspondence between the two patient samples (Table 3). Furthermore, these trends are very similar to the correlation matrix of clinician's ratings that can serve as the gold standard for these same symptom scales.[2] The results verify and support the internal structure of the TMJ Scale and further demonstrate that the cross-validation research sample well represents the general TMD clinical population. This supports placing a very high degree of confidence in the previous accuracy (sensitivity and specificity) and reliability results and in the clinical efficacy of this assessment tool.

Correlation studies on a large, national patient sample allow further inquiry into certain aspects of the clinical presentation of TMDs. For example, while joint dysfunction is correlated to a statistically significant extent ($P < .001$) with pain report ($r = .42$), palpation pain ($r = .30$), perceived malocclusion ($r = .31$), and range of motion limitation ($r = .57$), it is not highly correlated with psychological factors ($r = .10$) or stress ($r = .16$) (Table 3). The psychological factors scale correlates the highest with stress ($r = .68$) and chronicity ($r = .56$) and also has a strong and statistically significant ($P < .001$) association with non-TMD ($r = .51$), degree of pain ($r = .35$), and palpation pain ($r = .32$) (Table 3). Such correlations do not infer causality, but that these correlated symptom variables have clinically meaningful associations with one another. These associations, as studied through such techniques as those described, may lead to further elucidation of the clinical nature of TMDs and their appropriate treatment.

The availability of a national clinical sample of 10,549 patients allows certain epidemiologic questions to be approached. One interesting issue is the distribution of TMD symptom severity in the TMD patient population. One of the subscales on the TMJ Scale, the Global Scale, is composed of selected items from the five physical symptom subscales that include, pain, palpation pain, perceived malocclusion, joint dysfunction, and range of motion limitation.[2] The Global Scale was constructed as an overall screening scale to detect the presence of the category of problems called TMDs. This scale measures the overall symptom severity of the TMD. Its sensitivity and specificity, as established in the cross-validation research sample, are 84.2% and 80.3%, respectively.[2,9] Other studies of its reliability and predictive values have shown this scale to be psychometrically and clinically robust.[2,6,7,11–13,18]

The distribution of TMD symptom severity in the TMD patient population sample closely resembles a normal or Gaussian distribution (Fig 1).[20] In symmetrical, one-peaked distributions the mean and median are identical. In the case of the Global Scale the mean is 1.83 and the median is 1.85. This almost symmetrical, bell-shaped distribution suggests that the symptom severity of TMDs is normally distributed in the TMD patient population. Other studies using ratio level data to measure the overall symptom severity of TMDs are not available for comparison at this time.

Women with TMDs seem to manifest a higher level of severity of all physical and psychological symptoms relative to men (Table 4). This result indicates that women, on the average, report somewhat more severe symptoms in the areas of pain, palpation pain, perceived malocclusion, joint dysfunction, and range of motion limitation. Women also experience greater severity of symptoms related to nonTMDs, psychological factors, and stress (Table 4). A t test of the means in the two groups shows that this increased perception of severity for women is statistically significant to $P < .001$ (joint dysfunction $P < .034$) for all symptom areas except chronicity ($P = .317$).

The Global Scale measures the overall severity of the TMD. The average score for women is 1.87 compared to 1.59 for men (Table 4). This difference, which appears clinically significant, is also statistically significant to $P < .001$. Therefore, women with TMDs seem to experience a greater overall severity in symptoms of TMDs relative to men. The elevated symptom severity found in women may help explain the high ratio of women to men seeking treatment for TMD reported in most studies.

A higher percentage of women has clinically significant levels of pain report, palpation pain, and joint dysfunction as compared to men (Table 5). Men and women do not differ significantly in the

areas of perceived malocclusion and range of motion limitation (Table 5), and also do not differ significantly in the prevalence of non-TMDs. A higher percentage of male TMD patients have clinically significant problems involving psychological factors and stress, while women show a slightly higher prevalence than men in the area of chronicity (Table 5).

Pain levels show a small but statistically significant ($P = .001$) increase from the <30 years to the 30 to 39 years age group and then decrease again in the over 40 years age group to about the original level of severity (Table 6). The sample size is large, and even very small differences may be expected to yield statistically significant results, ie, small P values. The magnitude of change is small and does not appear clinically significant. The percentage of patients with clinically significant levels of pain shows a similar trend, starting at 85.6% in the under 30 years group, increasing to 86.6% in the 30 to 39 years age group, and then decreasing to 85% in the 40 years and older group (Table 5). Again, the differences do not appear to have much clinical relevance. Therefore, the different age groups seem to experience about the same severity levels and prevalence rate for TMD pain.

Palpation pain and perceived malocclusion (Tables 5 and 6) both show a small, yet statistically significant ($P < .001$) increase in severity as group age increases from less than 30 years to 30 to 39 years, and then this level of severity remains the same for ages 40 years and above. Again, the clinical relevance of this small trend is questionable. The percentage of patients with clinically significant palpation pain and perceived malocclusion shows a slight peak in the 30 to 39 years age group, with slightly lower percentages in younger and older age ranges.

Joint dysfunction and range of motion limitation reveal a completely different trend (Tables 5 and 6). These physical symptoms start at their most severe levels in the youngest age group and then demonstrate a statistically significant ($P < .001$) and clinically relevant trend of lower severity as group age increases to 30 to 39 years and 40 years and over (Table 6). There is a corresponding decline in the prevalence of clinically significant joint dysfunction and range of motion limitation as group age increases (Table 5). For example, the percentage of patients with clinically significant joint dysfunction starts at 79.5% in the less than 30 years age group and then decreases to 73.4% and 67.8% in the 30 to 39 years and the 40 years and over age groups, respectively. This is a most interesting finding, suggesting that the prevalences and symptom severities of both joint dysfunction and range of motion limitation are lower in older age groups. However, this does not necessarily suggest that joint dysfunction decreases as the time duration of a TMD increases. Age and time duration of TMD are separate variables, and the results of studies for different time durations of TMD produce somewhat different findings, as discussed below.

The severity of non-TMDs does not show a marked change in different age ranges, although their prevalence as clinically significant problems seems to peak slightly at 64.7% in the 30 to 39 years age group.

The severity of psychological factors shows a statistically significant ($P < .001$) but not clinically significant increase as group age increases from less than 30 years to 30 to 39 years, and then does not increase significantly in the 40 years and over age group (Table 6). The prevalence of clinically significant psychological factors increases from 42.9% in the under 30 years age group, to 49.0% in the 30 to 39 years age group, and it remains approximately the same at 49.3% in the 40 years and over age group. The severity of stress peaks in the 30 to 39 years age group, with the prevalence of clinically significant stress at 61.3% in this age group, as opposed to 53.0% and 58.3% in the younger and older age groups, respectively.

The severity of chronicity shows a statistically significant ($P < .001$) trend upward as group age advances, which might be expected (Table 6). The magnitude of increase appears to be clinically relevant when comparing the under 30 years to the 40 years and over age groups. This is paralleled by an increase in the prevalence of chronicity as group age increases. The prevalence is 68.5% in the youngest age group, 76.5% in the middle age group, and 79.7% in the oldest age group (Table 5). This result suggests that in older age groups there is a higher severity and prevalence of chronicity in the TMD patient population.

Finally, the overall severity of TMDs as measured by the Global Scale does not show a statistically significant change between the under 30 years and 30 to 39 years age groups (Table 6). In fact, the severity actually shows a statistically significant ($P < .001$) decline in the oldest age group relative to both of the younger age groups (Table 6). However, this decrease in severity is small and of questionable clinical significance.

A similar analysis of changes in severity and prevalence of symptoms may be undertaken with regard to the length of time TMDs have been present. Pain report, palpation pain, and range of

motion limitation all show similar trends as time duration increases (Tables 5 and 7). These symptom severities show a statistically significant ($P < .001$) increase from 1-5 months to 6 months-2 years, and then decrease slightly in the over 3 years' duration group. The changes found are all small and do not suggest an important clinical trend. Likewise, the percentage of patients with clinically significant problems in pain report and range of motion limitation generally shows a small increase in the 6 month to 2 year group over the 1 to 5 month group, and then a leveling off or decrease in the over 3 years' duration group (Table 5). Palpation pain does show a rather pronounced peak in prevalence in the 6 month to 2 years' duration group.

Perceived malocclusion and joint dysfunction both show a progressive and statistically significant ($P < .001$) increase in severity and prevalence as one moves from a time duration of 15 months to 6 months-2 years and then to over 3 years (Tables 5 and 7). The data suggest that the longer a TMD is present, the more severe and the more prevalent the perception of malocclusion and the symptoms associated with joint dysfunction. This is interesting since pain report and palpation pain do not increase in severity or prevalence when comparing the 6 month-2 year duration to the over 3 year duration group. Therefore, the greater severity of joint dysfunction symptoms found in groups having longer time durations of TMD is not associated with a corresponding increase in pain level. This result is consistent with the hypothesis that TMDs are slowly progressive, deteriorating processes. However, the present study is not a longitudinal study of patients over time, but rather a cross-sectional study of symptoms of joint dysfunction measured at one point in time within groups having a specified time duration of TMD. In other words, this study examines differences between groups, not changes over time. Prospective studies of changes in symptom severity over time in individual patients within specific diagnostic groups will help clarify these findings.

The severity of non-TMDs does not show a significant increase as time duration of TMDs increases (Table 7). However, the prevalence of non-TMDs increases from 58.0% in the 1 to 5 months group to 62.2% in the 6 months to 2 years group, and then increases further to 66.1 % in the over 3 years group. This result suggests that the longer a TMD is present, the higher the probability that a non-TMD will be concurrently present.

Psychological factors, stress, and chronicity all show the same increasing trends with increasing time duration of TMD (Tables 5 and 7). The severity of each of these shows a statistically significant ($P < .001$ to $P < .004$) increase with the time duration of TMD. The prevalence of each factor also shows an increase with time duration. The magnitude of changes in psychological factors is quite small and of questionable clinical significance. The overall trends suggest that the longer a TMD is present the more severe and the more prevalent are symptoms related to psychological factors, stress, and chronicity. Again this association does not prove causality. It is not possible, based on this research alone, to know whether the increased severity and prevalence of psychological factors, stress, and chronicity are caused by the TMD or vice versa. However, the association is clear. In the case of stress, 50.4% of patients having TMDs for 1 to 5 months have clinically significant levels of stress. This increases to 55.0% in the 6 months to 2 years group and increases further to 64.7% in the over 3 years group. This study indicates that the longer the TMD is present, the higher the severity and prevalence of stress experienced by the patient. Of course, the converse is true, that the higher the severity of stress, the longer a TMD is likely to be present. Again, causality cannot be determined based on this data alone.

The overall severity of TMDs, as measured by the Global Scale, increases to a statistically significant ($P < .001$) degree when moving from 1 to 5 months' duration to 6 months to 2 years' duration (Table 7). The increase in severity is not pronounced and is of marginal clinical significance. The severity does not change further in the longer time duration group of over 3 years. If one accepts the arbitrary cut off of 6 months as defining a chronic problem, then the overall severity of TMDs in the acute group (1 to 5 months) is somewhat lower than the overall severity in the chronic TMD group. Thus, patients in the chronic group appear symptomatically somewhat more impaired than those experiencing acute problems. The finding that patients with chronic problems are more symptomatically impaired is not inconsistent with the data showing that the symptom severity of TMDs is lower in older age groups. The age of the patient does not determine whether the patient has a chronic problem, ie, age and chronicity are independent characteristics.

Conclusion

The sensitivity, specificity, predictive values, and statistical reliability of the TMJ Scale were origi-

nally determined during cross-validation studies on large, research-based patient samples. It had been assumed that these research-based samples were selected in such a manner that their demographic characteristics and test responses would be representative of the clinical population in which the TMJ Scale would ultimately find use. The present study demonstrates that the test scores and various demographic variables including gender, age, marital status, education level, ethnic/racial distribution, and time duration of TMDs that characterize the original TMJ Scale cross-validation research sample accurately represent the general TMD patient population in which the TMJ Scale is now being used. Pearson correlation matrices in the original research sample and in the national clinical sample support the internal structural integrity of the TMJ Scale and demonstrate that the pattern of symptom variations in the two patient samples is very similar. This suggests a high degree of confidence in the results of previous accuracy (sensitivity and specificity) and reliability results and in the clinical efficacy of this assessment tool. The overall symptom severity of TMDs appears to be normally distributed in the TMD patient population. Epidemiologic results show important differences in both severity and prevalence of clinically significant symptoms when comparing men to women, different age ranges, and differing durations of TMDs. Women with TMDs report a higher level of severity of all physical and psychological symptoms relative to men. This may help explain the high female-to-male ratio in patients seeking treatment. However, a higher percentage of male TMD patients have clinically significant problems involving psychological factors and stress, while women show a slightly higher prevalence than men in the area of chronicity. Joint dysfunction and range of motion limitation symptoms show the highest severities and prevalence in the youngest age group and then demonstrate a continuing and statistically significant trend of lower severity as group age increases. Psychological factors and stress do not show a continuous increase in severity or prevalence in older age groups. The overall symptom severity of TMDs does not increase in older age groups. Pain, palpation pain, and range of motion limitation all peak in severity and prevalence in the 6 months to 2 years time duration of TMDs, while perceived malocclusion and joint dysfunction show higher levels of symptom severity and prevalence in groups having longer time durations of TMD. Therefore, the greater severity of joint dysfunction symptoms found in groups having longer time durations of TMD is not associated with a corresponding increase in pain level. Prospective studies of changes in symptom severity over time in individual patients within specific diagnostic groups will help clarify these findings. There is an association between increasing time duration of TMD and increasing severity and prevalence of psychological factors, stress, and chronicity. However, this research cannot distinguish whether longer time duration causes increased psychological problems, stress, and chronicity as opposed to the latter factors causing longer time duration of the TMDs. The data also suggest that patients having chronic problem (durations over 6 months) are symptomatically more impaired than those experiencing acute problems.

References

1. Griffiths R. Report of the President's conference on the examination, diagnosis, and management of temporomandibular disorders. J Am Dent Assoc 1983;106:75–77.
2. Levitt SR, Lundeen T, McKinney M. The TMJ Scale Manual. Durham, NC: Pain Resource Center, 1987.
3. Derogatis L. SCL 90 R: Administration, Scoring and Procedures Manual-II. Towson, MD: Clinical Psychometric Research, 1983.
4. Millon T, Green C, Meagher R. Millon Behavioral Health Inventory Manual. Minneapolis: National Computer Systems, 1979.
5. Lundeen TF, Levitt SR, McKinney MW. Evaluation of temporomandibular joint disorders by clinician ratings. J Prosthet Dent 1988;59:202–211.
6. Levitt SR, Lundeen TF, McKinney MW. Initial studies of a new assessment method for temporomandibular joint disorders. J Prosthet Dent 1986;59:490–495.
7. Lundeen TF, Levitt, SR, McKinney MW. TMJ Scale: Discriminative ability of the TMJ scale: age and gender differences. J Prosthet Dent 1986;56:84–92.
8. Lundeen TF, Levitt SR, McKinney MW. Clinical applications of the TMJ Scale. J Craniomand Pract 1988;6: 339–345.
9. Levitt SR, McKinney M, Lundeen T. The TMJ Scale: cross-validation and reliability studies. J Craniomand Pract 1988;6:17–25.
10. Lundeen T, Scruggs R, McKinney M, Daniel S, Levitt S.: TMD symptomatology among denture patients. J Craniomandib Disord Facial Oral Pain. 1990;4:40–45.
11. Levitt SR. The predictive value of the TMJ Scale in detecting psychological factors and non TM disorders in the patient with a temporomandibular disorder. J Craniomand Pract 1990;8:225–233.
12. Levitt SR. Predictive value: A model for dentists to evaluate the accuracy of diagnostic tests for temporomandibular disorders and applied to the TMJ Scale. J Prosthet Dent 1991;66:385–390.
13. Levitt SR. Predictive value of the TMJ Scale in detecting clinically significant symptoms of temporomandibular disorders. J Craniomandib Disord Facial Oral Pain 1990;4:177–185.

14. Spiegel EP, Levitt SR. Measuring symptom severity and treatment outcome of temporomandibular disorders with the TMJ Scale: Case report. J Craniomand Pract 1990;8: 353–358.
15. Spiegel EP, Levitt SR. Measuring symptom severity with the TMJ Scale. J Clin Orthod 1991;25:21–26.
16. Levitt SR, Spiegel EP, Claypoole WH. The TMJ Scale and undetected brain tumors in patients with temporomandibular disorders. J Craniomand Pract 1991;9:152–158.
17. Levitt S, McKinney M, Willis A. Measuring the impact of a dental practice on TM disorder symptoms. J Craniomand Pract 1993;11:211–216.
18. Pocock P, Mamandras A, Bellamy N. Evaluation of an anamnestic questionnaire as an instrument for investigating potential relationships between orthodontic therapy and temporomandibular disorders. Am J Orthod Dentofac Orthop 1992;102:239–243.
19. Nie N, Jenkins J, Bent D, Hull C, Steinbrenner K. SPSS-Statistical Package for the Social Sciences, ed 2. New York: McGraw-Hill, 1975.
20. McNemar Q. Psychological Statistics, ed 3. New York: John Wiley, 1962.

Resumen

Validación de la Escala de la Articulación Temporomandibular (ATM) en una muestra nacional de 10,000 pacientes — Características epidemiológicas y demográficas.

El presente estudio realizado en un consultorio clínico evaluó a más de 10,000 pacientes con desórdenes temporomandibulares (DTM). El estudio demuestra que los puntajes asignados al examen, las variables demográficas y los patrones de severidad de los síntomas que caracterizan la muestra investigativa original de la Escala ATM, representan verazmente la población general de pacientes que sufren de DTM y en quienes se está utilizando actualmente la Escala ATM. Los resultados indican que existe un alto grado de confianza en la eficacia clínica de este medio de evaluación. Se determinó que la severidad de los síntomas de los DTM en general, estaba distribuída normalmente en la población de pacientes. Las mujeres afectadas por los DTM reportan un grado de severidad más alto en cuanto a todos los síntomas psicológicos y físicos, en comparación a los hombres. Esto puede explicar el hecho de que proporcionalmente más mujeres buscan tratamiento en comparación a los hombres. Sin embargo, un porcentaje más alto de hombres con DTM tiene problemas psicológicos relacionados al stress clínicamente significativos, en comparación a las mujeres. La severidad y prevalencia de los síntomas asociados con la disfunción articular y la limitación en la extensión del movimiento son más bajos en las personas más viejas. Sin embargo, en general la severidad de los síntomas de los DTM no es más alta en las personas más viejas. No obstante, la severidad y prevalencia de los síntomas asociados con la disfunción articular son mayores en los grupos que han padecido de DTM por más tiempo, aunque los niveles de dolor no siguen esta tendencia. También existe una asociación entre la duración del desorden temporomandibular y la severidad de los problemas psicológicos y la cronicidad. Los pacientes con problemas crónicos están más impedidos sintomáticamente que aquellos con problemas agudos.

Zusammenfassung

Wertung der TMJ Scale in einer nationalen Stichprobe von 10,000 Patienten: Demographische und epidemiologische Gesichtspunkte.

In der vorliegenden Studie wurden über 10,000 Patienten bezüglich Myoarthropatien des Kausystems (MAP) klinisch untersucht. Die Resultate zeigten, dass die demographischen Variabeln, die Symptomsmuster und die Testergebnisse, die in der ursprünglich zur Entwicklung der TMJ Scale verwendeten Stichprobe gefunden wurden, genau denjenigen in einer MAP-Patienten-Population, in der die TMJ Scale heute üblicherweise verwendet wird, entsprechen. Die Resultate lassen auf einen hohen Grad an Zuverlässigkeit dieser klinischen Untersuchungsmethode schliessen. Bezüglich Schweregrad waren die MAP-Symptome in der untersuchten Patientenpopulation normalverteilt. Frauen mit MAP zeigen eine gegenüber Männern erhöhte Schwere aller auftretenden physischen und psychischen Symptome. Dies mag erklären, warum Frauen häufiger als Männer eine Therapie wünschen. Klinisch signifikante psychologische und stressbezogene Probleme finden sich zu einem höheren Prozentsatz bei Männern als bei Frauen mit MAP. Gelenksymptome und Beweglichkeitseinschränkungen kommen weniger oft und weniger ausgeprägt in älteren Patientengruppen vor; der gesamte Schweregrad der MAP-Symptome ist nicht grösser in älteren Gruppen. Schweregrad und Prävalenz der Symptome im Zusammenhang mit MAP sind grösser in Gruppen, in denen die MAP schon länger vorliegt, die Schmerzintensität folgt dieser Tendenz allerdings nicht. Ebenso existiert eine Assoziation zwischen der zeitlichen Dauer der MAP und Schweregrad und Chronizität der psychologischen Probleme. Patienten mit chronischen Problemen sind durch ihre Symptome stärker beeinträchtigt als solche mit akuten Problemen.