Predictors of Long-Term Temporomandibular Disorder Pain Intensity: An 8-Year Cohort Study

Flavia P. Kapos, DDS, MS
Senior Fellow Trainee
Department of Oral Health Sciences
School of Dentistry
University of Washington;
PhD Student
Department of Epidemiology
School of Public Health
University of Washington
Seattle, Washington, USA

John O. Look, DDS, MPH, PhD
Senior Research Associate
Division of TMD and Orofacial Pain
Department of Diagnostic and Biological Sciences
School of Dentistry
University of Minnesota
Minneapolis, Minnesota, USA

Lei Zhang, MS
Research Fellow
Biostatistical Design and Analysis Center (BDAC)
Clinical and Translational Science Institute
University of Minnesota
Minneapolis, Minnesota, USA

James S. Hodges, PhD
Professor
Division of Biostatistics
School of Public Health
University of Minnesota
Minneapolis, Minnesota, USA

Eric L. Schiffman, DDS, MS
Professor
Division of TMD and Orofacial Pain
Department of Diagnostic and Biological Sciences
School of Dentistry
University of Minnesota
Minneapolis, Minnesota, USA

Correspondence to:
Dr Flavia P. Kapos
University of Washington
1959 NE Pacific St, Box 357475
Seattle, WA 98195, USA
Email: kapos@uw.edu

Aims: To investigate, in individuals with pain-related temporomandibular disorder (TMD), the association of long-term pain intensity with baseline health-related quality of life (HRQoL) and jaw functional limitation. **Methods**: Of 513 cases with baseline pain-related TMD (masticatory muscle and/or temporomandibular joint [TMJ] pain), 273 were reevaluated after 8 years, and 258 of them had complete baseline data for Jaw Functional Limitation Scale (JFLS) scores and HRQoL measured by the Physical Component Summary (PCS) and Mental Component Summary (MCS) scores of the 12-item Short Form Health Survey and follow-up data for Characteristic Pain Intensity (CPI) from the Graded Chronic Pain Scale. Secondary analyses of existing data quantified the effects of primary (PCS, MCS) and secondary (JFLS) predictors on follow-up CPI by using multivariable linear regression. Sensitivity analyses considered differences between the included participants (n = 258) and those who were not included (n = 255) by using inverse probability weighting. Interactions of baseline predictors with age, sex, and baseline CPI were evaluated using multivariable linear regression. **Results:** The score for baseline PCS, but not MCS or JFLS, was associated with follow-up CPI (P = .012). One standard deviation (SD = 9.0)–higher baseline PCS score predicted an overall 3.2-point–lower follow-up CPI (95% confidence interval –5.8 to –0.7) after adjusting for age, sex, MCS, JFLS, and baseline CPI scores. However, the effect of PCS score was not uniform: the association between PCS and follow-up CPI scores was statistically significant for participants with baseline CPI ≥ 51.3/100 and clinically significant for participants with baseline CPI ≥ 68.7/100. Adjustment for TMD treatments and sensitivity analyses had negligible effect. **Conclusion:** In participants with moderate to severe baseline TMD pain intensity, higher baseline physical HRQoL predicted lower TMD pain intensity at 8 years follow-up. PCS score could contribute to a multifactorial long-term TMD pain prediction model.
J Oral Facial Pain Headache 2018;32:113–122. doi: 10.11607/ofph.1819

Keywords: *chronic pain, epidemiology, facial pain, quality of life, temporomandibular disorders*

Temporomandibular disorders (TMD) affect the masticatory muscles, temporomandibular joints (TMJs), and associated structures.[1] Pain, the most frequent reason patients seek TMD treatment, affects approximately 10% of adults.[2] It has been reported that most TMD cases tend to remit or present as recurrent pain episodes. Approximately 15% of patients who seek care progress to chronic TMD pain,[3] often defined as pain lasting for at least 3[4] to 6 months.[5] Causal factors for persistent TMD pain are not clear, although psychosocial factors differ significantly in TMD cases compared to pain-free controls[6–9] and are associated with chronic TMD pain.[10–12] Previous studies have suggested that biopsychosocial factors, including poor general health, can predict onset of chronic pain conditions such as widespread pain[13] and musculoskeletal disorders,[14–18] including TMD.[19–22]

Health-related quality of life (HRQoL) measures have been used to assess an individual's functioning and disease burden for many chronic physical conditions, including headaches,[23] arthritis,[24,25] back pain,[16,24,25] and TMD.[7,26,27] The 12-item Short Form Health Survey (SF-12)[28,29] is a commonly used HRQoL questionnaire derived from the 36-item version (SF-36).[30] The SF-12 and SF-36 evaluate eight health domains

(physical functioning, role physical, bodily pain, general health, mental health, role emotional, social functioning, and vitality), yielding two summary scores: the Physical Component Summary (PCS) and the Mental Component Summary (MCS). SF-12 and SF-36 scores have been shown to differ significantly between healthy controls and patients with migraine,[31,32] fibromyalgia,[32] and TMD,[26] and the SF-12 score has been found to have a dose-response relationship with severity of various chronic health conditions.[33] In a multivariable model, the SF-12 bodily pain and general health subscales were among the 11 best predictors for new-onset TMD out of 202 putative risk factors evaluated in a large cohort.[34] To date, no studies have evaluated the capacity of the SF-12 scores to predict long-term pain intensity in TMD participants.

According to the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT),[35] there are two broad areas of assessment for physical functioning and HRQoL: generic measures such as the SF-12 and disease-specific physical functioning measures such as the Jaw Functional Limitation Scale (JFLS).[36] In a large case-control study, the JFLS score was significantly worse in TMD cases compared to pain-free controls.[37]

Although some factors have been identified as having cross-sectional and retrospective associations with TMD pain or as longitudinal risk factors for new-onset TMD, factors that can identify existing TMD cases at risk for long-term TMD pain are largely unknown. The present study aimed to address this gap and elucidate the effects of a few potential predictors of TMD pain intensity. A unique opportunity to do so was available in conducting secondary data analyses of a well-defined cohort of TMD pain cases from the multicenter Validation Project[38] and its 8-year follow-up, the TMJ Impact Project.[39] The present study's aim was to investigate, in individuals with pain-related TMD, the association of long-term TMD pain intensity with baseline HRQoL and jaw functional limitation. This aim was addressed by quantifying the effect of the baseline PCS, MCS, and JFLS scores on follow-up TMD pain intensity, measured by the Characteristic Pain Intensity (CPI)[5] score from the Graded Chronic Pain Scale (GCPS).[5] It was hypothesized that baseline PCS and MCS would be negatively associated with long-term TMD pain intensity and that baseline JFLS score would be positively associated.

Materials and Methods

Study Sample

At baseline, participants in the Validation Project were recruited at the University of Minnesota, University of Washington, and University at Buffalo between 2003 and 2006. Participants either were referred from local health care providers to each university-based TMD clinic or they responded to community advertisements. Institutional review board approval was obtained at each study site before study initiation, and all participants provided informed consent. Enrollment was consecutive until two-thirds of the Validation Project's target recruitment[38] was achieved; thereafter, participants with less common TMD diagnoses and of older age were selectively enrolled. Thus, the Validation Project participant population was a convenience sample of individuals 18 to 70 years old, consisting of healthy controls and clinical and community TMD cases with the full spectrum of TMD signs and symptoms but without significant non-TMD pain comorbidities.

Figure 1 shows the flow of study participants from the baseline Validation Project to the follow-up TMJ Impact Project and, ultimately, the subset of participants eligible for the present study's secondary data analyses. The Validation Project included 513 participants who at baseline had a pain-related gold-standard TMD diagnosis (masticatory muscle and/or TMJ pain) and who were therefore potentially eligible for the present study. Approximately 8 years later, funding was approved for the TMJ Impact Project to follow up 400 of the 705 total Validation Project participants, with the sampling frame limited to those who had previously agreed to be contacted. Actual follow-up of the TMJ Impact Project included 401 participants, 273 who had pain-related TMD diagnoses at baseline and 128 who had nonpainful TMD diagnoses or were pain-free controls. Among the 304 Validation Project participants not followed up in the TMJ Impact Project, 240 had pain-related TMD diagnoses and 64 had nonpainful TMD diagnoses or were pain-free controls.

Inclusion in the present study required complete data for: *(1)* at least one baseline gold-standard pain-related TMD diagnosis (masticatory muscle and/or TMJ pain); *(2)* baseline CPI score greater than 0; *(3)* baseline SF-12 (PCS and MCS) scores; and *(4)* follow-up CPI score. Those with pain-related TMD diagnoses who were not included in this study (n = 255) consisted of those not followed up in the TMJ Impact Project (n = 240) and those excluded from the present analysis (n = 15) for the following reasons: missing data for baseline SF-12 (n = 8); missing data for follow-up CPI (n = 1); and baseline pain-related TMD diagnoses with a CPI of 0 (n = 6). Thus, the present study included 258 participants with TMD pain at baseline.

For the purpose of this study, TMD pain at baseline was defined as the presence of at least one pain-related TMD diagnosis and a CPI score greater than 0. Long-term TMD pain was defined by its

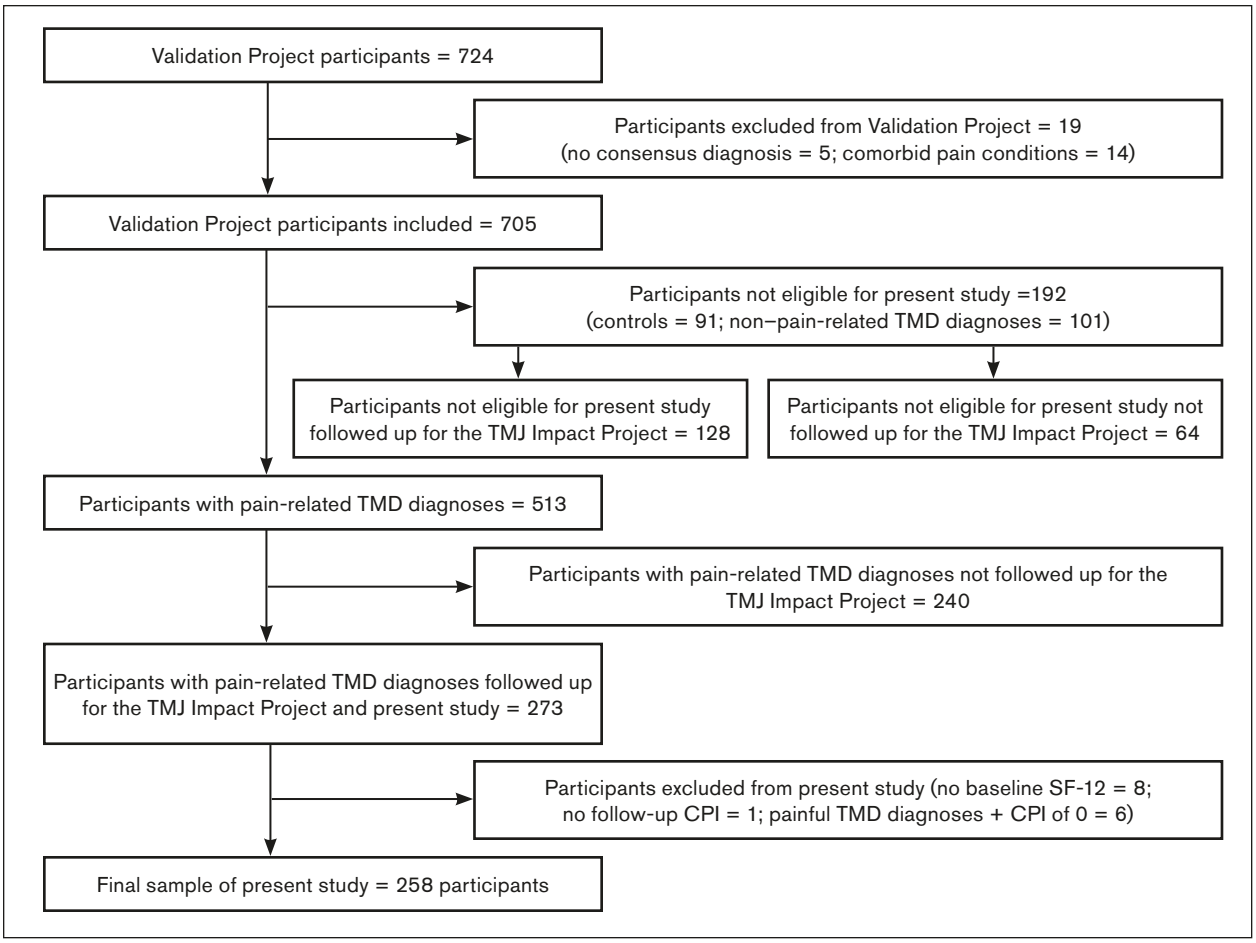

Fig 1 Flow diagram of study participants. CPI = Characteristic Pain Intensity.

presence at both baseline and follow-up; ie, by evidence that conditions causing the TMD pain had remained unresolved for a longer time than would normally be expected.[3,4] TMD pain between baseline and follow-up visits was not assessed; whether it was continuous or episodic was not known.

Clinical Assessment and TMD Diagnoses

Baseline pain-related TMD diagnoses were established by consensus between two clinical examiners at each study site after each independently assessed the participants using a comprehensive history and clinical examination protocol.[38] The Validation Project's complete methods and eligibility criteria have been previously reported.[38] At follow-up, each participant was seen at the same study site by only one of the examiners. Pain-related TMD diagnoses at follow-up were algorithmically derived from history and examination data collected using the Axis I Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) protocol.[40] At all three study sites, the examiners at baseline and at follow-up were the same TMD experts. Inter-rater reliability of examiners, assessed with kappa, was 0.83 for masticatory muscle pain and 0.85 for TMJ pain at baseline[41] and 0.84 for masticatory muscle pain and 0.76 for TMJ pain at follow-up (unpublished data).

Outcome Measure

The study's primary outcome measure was follow-up TMD pain intensity, measured by the CPI score from the GCPS.[5] The CPI score ranges from 0 to 100 and is calculated from three items of self-reported TMD pain intensity: *(1)* pain at present time; *(2)* worst pain in the last 6 months; and *(3)* average pain in the last 6 months. Each item consists of a 0–10 numeric rating scale (NRS), where 0 indicates no pain and 10 indicates pain as bad as it could be. The 3 scores are averaged and multiplied by 10 to obtain the CPI score.

Baseline Predictors

Baseline questionnaires included the SF-12 version 2[28] and the JFLS-20.[36,42] The SF-12 is a reliable and valid HRQoL questionnaire[29,33,43] used in both clinical and research settings. The SF-12's concurrent validity with the SF-36 has been established for PCS and MCS[29,33] and with the EuroQol 5-Dimension (EQ-5D), another HRQoL questionnaire.[44] MCS and

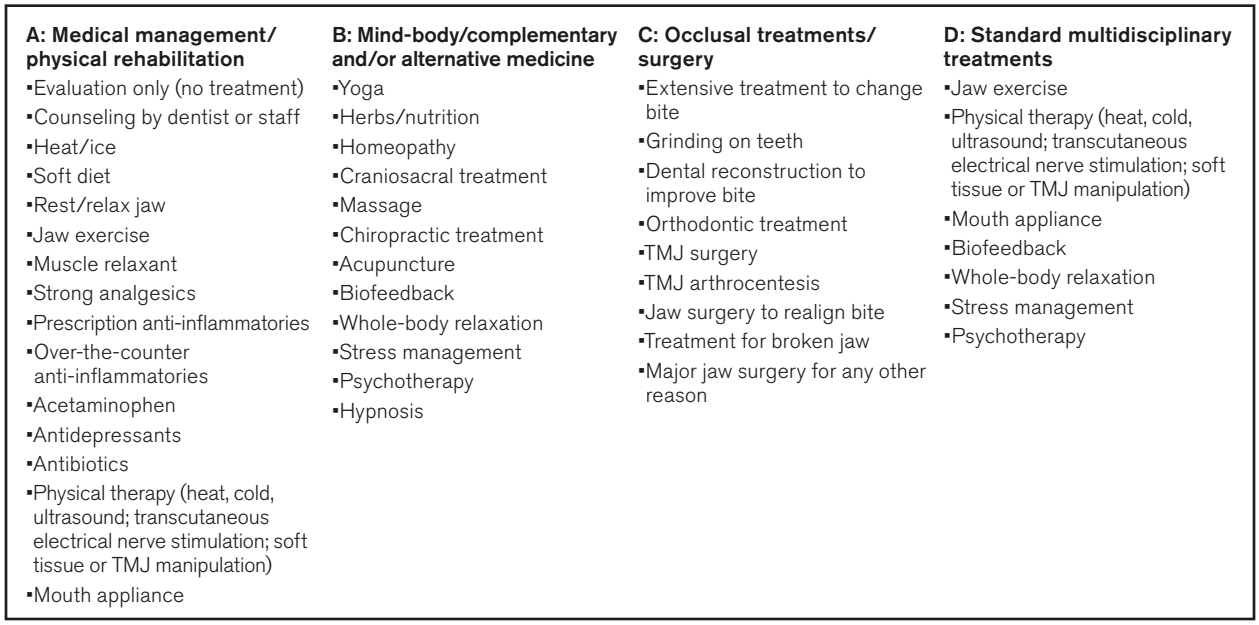

Fig 2 Survey of treatments received. These items and treatment categories were used for multivariable analyses adjustments.

PCS summary scores range from 0 to 100, with a higher score indicating better quality of life. The JFLS consists of 20 items, each with a 0–10 NRS, where 0 indicates no jaw limitation and 10 severe jaw limitation for each specified activity. The JFLS has three subscales: mastication, vertical jaw mobility, and verbal and emotional expression; their average yields a global jaw functional limitation score on a 0–10 scale. The JFLS has been validated in TMD patients with reliability coefficients of 0.82 for persons and 0.99 for items.[42]

Participant self-reports of TMD treatments during the follow-up period consisted of a questionnaire with 39 yes/no items, which were clustered into three groups for analyses:

- A: Medical management/physical rehabilitation
- B: Mind-body/complementary and/or alternative medicine
- C: Occlusal treatments/surgery

A fourth grouping (D: standard multidisciplinary treatments) was formed from a subset of 10 of these items (Fig 2). Three definitions for TMD treatments received were considered as adjusters in the analyses: *(1)* any treatment (yes/no); *(2)* any Group A (yes/no), Group B (yes/no), or Group C (yes/no) treatments; and *(3)* any Group D treatments (yes/no).

Statistical Analyses

Linear regression was used to investigate baseline PCS, MCS, and JFLS scores as predictors of follow-up CPI in multivariable models adjusted for baseline characteristics: age, sex, TMD treatments received, and baseline CPI. Linear regression coefficient estimates are often reported as the difference in outcome associated with a one-unit increase in a given predictor. Since the present study's multiple study predictors were measured on different scales, the respective effects of one-unit increases were not directly comparable; thus, for each predictor, the difference in the outcome associated with one standard deviation (SD) increase in the predictor was computed. SD standardization of coefficient estimates is often employed in epidemiology to facilitate direct comparison of the effect sizes of study predictors.

Additional Analyses. Since the analyzed subset of cases differed from excluded cases in average age and global JFLS score, the analyzed sample was differentially weighted to adjust for these differences by using the inverse of each person's probability of being an included case.[45] These probabilities were estimated in a secondary logistic regression analysis with the outcomes yes (included case) or no (excluded case). Covariates for the weighting analysis were age and baseline JFLS.

Given that about half of the TMD pain participants from the Validation Project were not included in the analysis in the interest of generalizing the present study's findings to all 513 TMD pain cases in the Validation Project, the authors did a series of variant analyses, adding to the main analysis interactions of the main effects with selected population characteristics: sex, age (continuous), or age (categorical).

To determine whether the association between baseline predictors and follow-up CPI differed according to baseline CPI, the authors also estimated and tested their interactions with baseline CPI.

Table 1 Characteristics of Participants With and Without Follow-up Pain-Related TMD

			Follow-up pain-related TMD diagnoses	
Variable	Category	Overall N = 258	Yes n = 186	No n = 72
Sex, n (%)	Male	31 (12.0)	18 (9.7)	13 (18.1)
	Female	227 (88.0)	168 (90.3)	59 (81.9)
Age (y)	Mean (SD)	37.8 (13.0)	38.4 (13.2)	36.2 (12.7)
	(Min, Max)	(18.0, 67.0)	(18.0, 67.0)	(18.0, 65.0)
Baseline CPI score (0–100)	Mean (SD)	50.2 (20.0)	51.0 (19.5)	48.1 (21.4)
	(Min, Max)	(6.7, 100.0)	(10.0, 93.3)	(6.7, 100.0)
Baseline SF-12 PCS score (0–100)	Mean (SD)	50.6 (9.0)	49.8 (9.5)	52.6 (7.0)
	(Min, Max)	(15.7, 65.4)	(24.1, 65.4)	(15.7, 64.5)
Baseline SF-12 MCS score (0–100)	Mean (SD)	49.6 (9.1)	49.1 (9.21)	50.8 (8.7)
	(Min, Max)	(22.8, 67.3)	(22.8, 65.2)	(23.3, 67.3)
Baseline TMD pain location, n (%)	TMJ pain only	35 (13.6)	26 (14.0)	9 (12.5)
	Muscle pain only	3 (1.2)	1 (0.5)	2 (2.8)
	Both	220 (85.3)	159 (85.5)	61 (84.7)
Baseline JFLS global score (0–10)	Mean (SD)	1.8 (1.4)	1.9 (1.4)	1.6 (1.5)
	(Min, Max)	(0.0, 8.0)	(0.0, 6.5)	(0.00, 8.0)
Follow-up CPI (0–100)	Mean (SD)	28.1 (19.9)	34.5 (18.0)	11.6 (14.2)
	(Min, Max)	(0.0, 86.7)	(6.7, 86.7)	(0.0, 66.7)
Follow-up TMD pain location, n (%)	None	72 (27.9)	0	72 (100.0)
	TMJ pain only	35 (13.6)	35 (18.8)	0
	Muscle pain only	5 (1.9)	5 (2.7)	0
	Both	146 (56.6)	146 (78.5)	0

CPI = Characteristic Pain Intensity; PCS = Physical Component Summary; MCS = Mental Component Summary; JFLS = Jaw Functional Limitation Scale; SD = standard deviation.

Statistical significance was set at $P \leq .025$ for each of the two primary analyses (baseline PCS and MCS scores predicting follow-up CPI [ie, Bonferroni correction]). For all other tests considered exploratory, statistical significance was defined as $P < .05$. All analyses were done using SAS (v. 9.3; SAS Institute). Graphs were plotted in R (http://www.r-project.org).

Table 2 Difference in Follow-up Characteristic Pain Intensity (CPI) Predicted by One Standard Deviation (SD)–Higher Baseline Predictors

Baseline predictor	SD coefficient estimates	95% CI	*P* value
PCS (SD 9.0)	–3.2	(–5.8, –0.7)	.012
MCS (SD 9.1)	–1.2	(–3.6, 1.1)	.30
JFLS (SD 1.4)	0.4	(–2.2, 3.0)	.78

SD coefficient estimates represent β (regression line slope), interpreted as the difference in follow-up CPI associated with one SD–higher baseline predictors when controlling for age and sex, the other baseline predictors in the model, and baseline CPI. Please compare to Table 5, as the PCS estimate was not uniform for all levels of baseline CPI. PCS = Physical Component Summary; MCS = Mental Component Summary; JFLS = Jaw Functional Limitation Scale; CI = Confidence Interval.

Results

Elapsed time between the baseline and follow-up examinations averaged 8.0 ± 0.7 years (range 6.3 to 10.0). Of the 258 participants included in the analysis with TMD pain at baseline, 186 (72%) were also diagnosed with TMD pain at follow-up. Of the total 258 participants, 88% were women with a mean age of 38 ± 13 years (range 18 to 67). Table 1 presents characteristics of participants with and without follow-up pain-related TMD diagnoses. Considering all 258 participants, follow-up TMD pain intensity (CPI score) had an overall mean of 28.1 ± 19.9 points (range 0 to 86.7) on a 0–100 scale, an average of 22.1 points (43.8%) lower compared to baseline CPI scores (mean 50.2 ± 20.0 points). For the subset of participants with pain-related TMD diagnoses at follow-up, the mean follow-up CPI was 34.5 ± 18.0 points (range 6.7 to 86.7); that is, lower than baseline by an average of 15.7 points (31.3%). At follow-up, some participants reported a CPI score greater than 0 for a 6-month time frame but were classified as not having a pain-related TMD diagnosis. This situation was clinically possible because the diagnostic criteria were based on pain in the last month; hence, Table 1 includes some participants with follow-up CPI > 0 in the group without follow-up pain-related TMD diagnoses.

Multivariable Linear Regression

The multivariable linear regression model included all three baseline study predictors (PCS, MCS, and JFLS scores), as well as potential confounders (age, sex, and baseline CPI) (Table 2). One SD (9.0)–higher baseline PCS score predicted a 3.2-point (6.4%)–lower follow-up CPI (95% confidence interval [CI] –5.8 to –0.7,

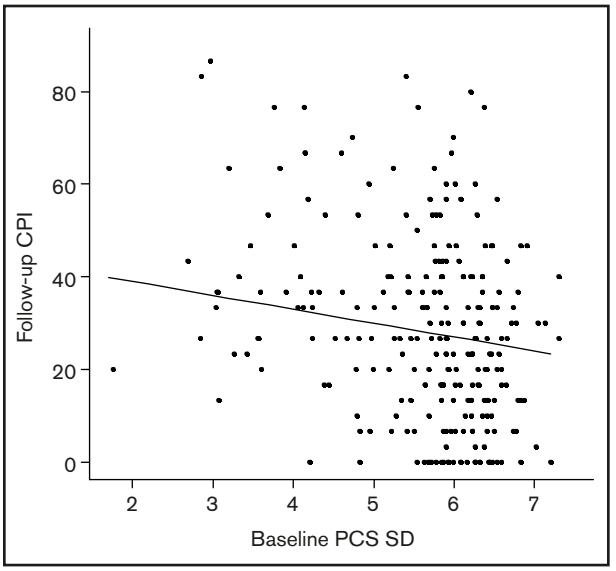

Fig 3 Linear association between baseline Physical Component Summary (PCS) and follow-up Characteristic Pain Intensity (CPI). One standard deviation (9.0)–higher baseline PCS was associated with a 3.2-point lower follow-up CPI score (95% confidence interval [CI] −5.5, −0.6; P = .012) after adjusting for age, sex, MCS, JFLS, and baseline CPI.

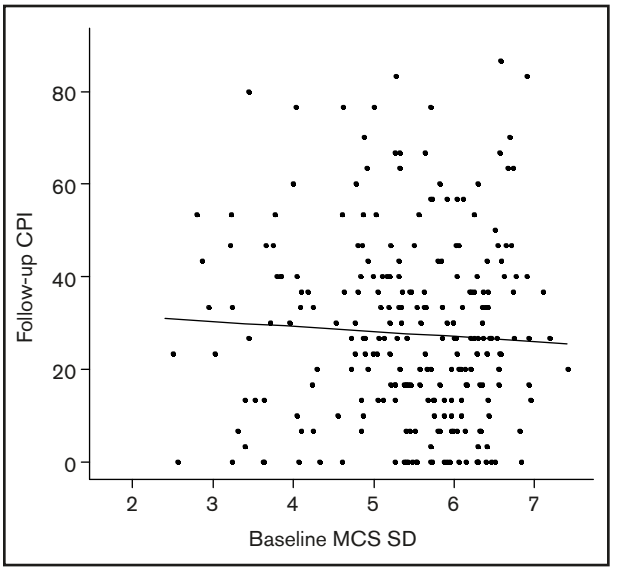

Fig 4 Linear association between baseline Mental Component Summary (MCS) score and follow-up Characteristic Pain Intensity (CPI). One standard deviation (9.1)–higher increase in baseline MCS predicted a 1.2-point lower mean follow-up CPI score; there was no significant linear association (95% CI −3.6 to 1.1, P = .30) after adjusting for age, sex, PCS, JFLS, and baseline CPI scores.

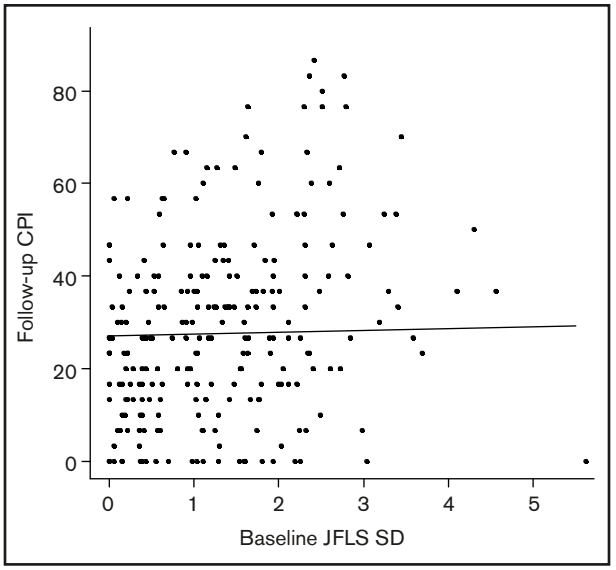

Fig 5 Linear association between baseline Jaw Functional Limitation Scale (JFLS) score and follow-up Characteristic Pain Intensity (CPI). One standard deviation (1.4)–higher baseline JFLS score predicted a 0.37-point higher mean follow-up CPI score; there was no significant linear association (95% CI −2.2 to 3.0, P = .78) after adjusting for age, sex, PCS, MCS, and baseline CPI scores.

P = .012) after adjusting for age, sex, MCS, JFLS, and baseline CPI (Fig 3). One SD (9.1)–increase in baseline MCS score predicted a 1.2-point (2.4%)–lower mean follow-up CPI (95% CI −3.6 to 1.1, P = .30) after adjusting for age, sex, PCS, JFLS, and baseline CPI scores (Fig 4). One SD (1.4)–higher baseline JFLS score predicted a 0.37-point (0.74%)–higher mean follow-up CPI (95% CI −2.2 to 3.0, P = .78) after adjusting for age, sex, PCS, MCS, and baseline CPI scores (Fig 5). After adjusting further for treatments received using each of the three ways of summarizing the treatments received, each predictor maintained similar statistical significance and overall magnitude of its effect (Table 3).

Sensitivity Analysis: Potential Bias due to Partial Follow-up

Compared to those cases who were not followed up or who were excluded from the analysis, cases included in the analysis were on average 2.6 years older (included = 37.8 ± 13.0 years, 95% CI 36.2 to 39.4; not included = 35.2 ± 13.1 years, 95% CI 33.6 to 36.8; P = .028). Included cases had a 0.39-point–lower average JFLS score at baseline (included = 1.8 ± 1.4 points, 95% CI 1.7 to 2.0; not included = 2.2 ± 1.5 points, 95% CI 2.0 to 2.4; P = .004). Although this baseline difference in JFLS score was less than 5% of the range for JFLS score (0.00 to 7.98), the statistical influence of this difference was investigated to rule out a systematic difference between cases included and not included in the analysis. A sensitivity analysis compared the unweighted linear regression estimates above with suitably weighted analyses in a series of variant regression models with and without adjusters (Table 4). Weighting created slight differences in the SD standardized coefficient estimates for predicting

Table 3 Difference in Follow-up Characteristic Pain Intensity (CPI) Predicted by One Standard Deviation (SD)–Higher Baseline Predictors Adjusted for Age, Sex, Baseline CPI, and TMD Treatments Received

	Any treatment		Any Group A, B, C		Any Group D	
Baseline predictor	SD coefficient estimates (95% CI)	*P* value	SD coefficient estimates (95% CI)	*P* value	SD coefficient estimates (95% CI)	*P* value
PCS (SD 9.0)	–3.0 (–5.5, –0.6)	.014	–3.6 (–6.0, –1.2)	.004	–3.2 (–5.8, –0.7)	.012
MCS (SD 9.1)	–1.2 (–3.4, 1.1)	.31	–0.81 (–3.0, 1.4)	.48	–1.2 (–3.5, 1.2)	.34
JFLS (SD 1.4)	0.47 (–3.0, 2.1)	.72	–0.31 (–2.8, 2.2)	.81	0.26 (–2.3, 2.9)	.84

SD coefficient estimates represent β (regression line slope), interpreted as the difference in follow-up CPI associated with one SD–higher baseline predictors when controlling for age, sex, the other predictors in the model, baseline CPI, and for treatments received, with the latter defined in three different ways: *(1)* any treatment; *(2)* any Group A, B, or C treatments; and *(3)* any Group D treatment. PCS = Physical Component Summary; MCS = Mental Component Summary; JFLS = Jaw Functional Limitation Scale; CI = Confidence Interval.

Table 4 Sensitivity Analyses for Study Predictors of Follow-up Characteristic Pain Intensity (CPI)

Type of analysis	Baseline predictors	Unweighted SD coefficient estimates (95% CI)	*P* value	Weighted SD coefficient estimates (95%CI)	*P* value
Unadjusted	PCS (SD 9.0)	–6.0 (–8.3, –3.6)	< .001	–6.1 (–8.5, –3.7)	< .001
	MCS (SD 9.1)	–0.8 (–3.3, 1.6)	.51	–0.5 (–3.0, 2.0)	.72
	JFLS (SD 1.4)	5.7 (3.3, 8.1)	< .001	5.1 (2.7, 7.4)	< .001
Adjusted for sex and age	PCS (SD 9.0)	–4.9 (–7.3, –2.4)	< .001	–5.2 (–7.8, –2.7)	< .001
	MCS (SD 9.1)	–1.4 (–3.7, 1.0)	.26	–1.0 (–3.4, 1.5)	.44
	JFLS (SD 1.4)	5.0 (2.6, 7.3)	< .001	4.5 (2.2, 6.8)	< .001
Adjusted for baseline PCS, MCS, JFLS, sex, age, baseline CPI, and any treatment	PCS (SD 9.0)	–3.0 (–5.5, –0.6)	.014	–3.4 (–6.0, –0.9)	.009
	MCS (SD 9.1)	–1.2 (–3.4, 1.1)	.31	–1.1 (–3.4, 1.2)	.34
	JFLS (SD 1.4)	0.5 (–3.0, 2.1)	.72	–0.7 (–3.2, 1.8)	.59

SD coefficient estimates represent β (regression line slope), interpreted as the difference in follow-up CPI associated with one SD–higher baseline predictors. Weighted analyses were based on the inverse probability weighting method to address bias associated with loss to follow-up. PCS = Physical Component Summary; MCS = Mental Component Summary; JFLS = Jaw Functional Limitation Scale; CI = Confidence Interval.

follow-up CPI, but the *P* values for these estimates were similar. No other baseline characteristics (sex, CPI, PCS, MCS, or pain location) differed significantly between included and excluded cases.

Generalization of Results to Population Subgroups

To assess the generalizability of the present study's findings to population subgroups, potential interactions were investigated between baseline predictors and three principal population characteristics. Interactions of sex with the three predictors showed *P* values ranging from .23 to .86. For age treated as a continuous measure, the range of *P* values was .25 to .81. For the age category of 18 to 35 years, the *P* value range was .11 to .48. Interactions with the category of 36 to 50 years showed *P* values ranging from .29 to .91.

Differences in the Association Between Baseline Predictors and Follow-Up TMD Pain Intensity According to Baseline TMD Pain Intensity

The multivariable linear regression estimate for the overall PCS effect size associated with follow-up CPI was an average 3.2-point–lower follow-up CPI score for every one SD increase in baseline PCS score. Although this estimate was statistically significant (*P* = .012), an analysis including an interaction of baseline CPI with baseline PCS score found that the interaction was statistically significant (*P* = .008), implying that the PCS effect size associated with follow-up CPI depended on the baseline CPI. In particular, baseline PCS score predicted significantly larger differences in follow-up CPI in participants who had higher baseline CPI. For every SD (20.0)–higher baseline CPI, the follow-up CPI predicted by one SD–higher baseline PCS score was 3.14 points lower (95% CI 0.82 to 5.47). Using this analysis, Table 5 presents examples of the estimated effect of baseline PCS score on follow-up CPI for four values of baseline CPI: *(1)* minimum observed; *(2)* maximum observed; *(3)* minimum baseline CPI for which the association between baseline PCS score and follow-up CPI was statistically significant; and *(4)* minimum baseline CPI for which the association between baseline PCS score and follow-up CPI was clinically significant (at least 10% change from mean baseline CPI, or 5.0 points) (Fig 6). This figure shows that the effect of PCS score was statistically significant only for participants with baseline CPI > 51.3 and clinically significant only for participants with baseline

Table 5 Difference in Follow-up Characteristic Pain Intensity (CPI) Predicted by One SD (9.0)–Higher Baseline Physical Component Summary (PCS) According to Baseline CPI

Baseline CPI	SD coefficient estimates	95% CI	*P* value
Minimum (6.7)	3.7	(–2.6, 10.1)	.25
Maximum (100)	–9.4	(–15.2, –3.7)	.001
Statistical significance (51.3)	–2.6	(–5.1, 0.0)	.050
Clinical significance (68.7)	–5.0	(–7.9, –2.1)	.001

SD coefficient estimates represent β (regression line slope), interpreted as the difference in follow-up CPI associated with one SD (9.0)–higher baseline PCS score at different values of baseline CPI, controlling for age, sex, MCS, and JFLS scores. MCS = Mental Component Summary; JFLS = Jaw Functional Limitation Scale; CI = Confidence Interval.

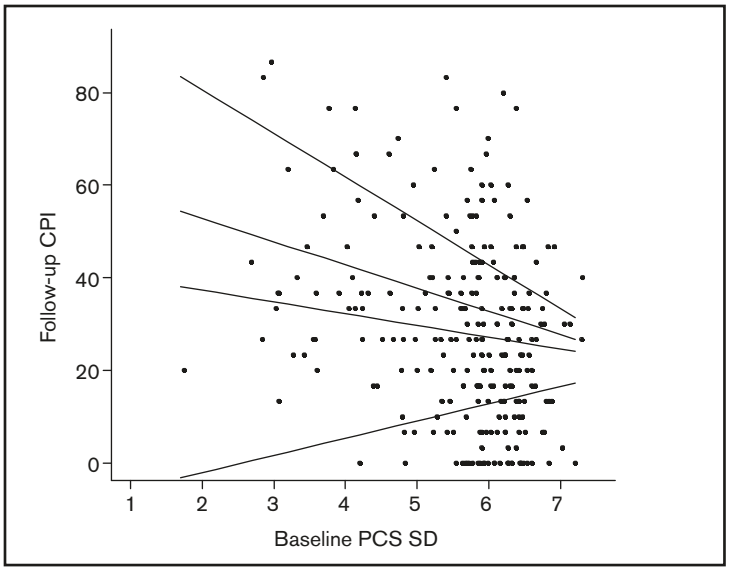

Fig 6 Linear association between baseline Physical Component Summary (PCS) score and follow-up Characteristic Pain Intensity (CPI) according to baseline CPI. After adjusting for age, sex, MCS, and JFLS score at four values of baseline CPI (in ascending order of depicted lines): [a]minimum baseline CPI observed; [b]minimum baseline CPI for statistical significance of the association between baseline PCS score and follow-up CPI; [c]minimum baseline CPI for clinical significance (at least 10% of mean baseline CPI, or 5.0 points) of the association between baseline PCS score and follow-up CPI; and [d]maximum baseline CPI observed.

CPI > 68.7. Conversely, baseline PCS score had no significant effect on follow-up CPI for about half of the participants, those with baseline CPI < 51. The association between baseline MCS and JFLS scores with follow-up CPI did not depend on baseline CPI (*P* = .24 and *P* = .81 for the interactions, respectively).

Discussion

To the authors' knowledge, this is the first study to report the effect size and clinical significance of PCS, MCS, and JFLS scores as predictors of long-term TMD pain intensity in individuals with baseline TMD pain. The present study found that PCS had a statistically significant association with follow-up CPI and that this association was also clinically significant for participants with higher baseline TMD pain intensity. The differences in outcome predicted by MCS and JFLS scores were not statistically significant.

Clinical Significance of Differences in Pain Intensity

Statistically significant outcomes may not attain clinical significance. IMMPACT has recommended that a decrease of 10% to 20% in self-reported pain intensity is minimally important improvement, 30% or greater is moderately important, and 50% is substantially important.[46] Also, a difference of 2 points on a 0–10 NRS or a 30% change from baseline was considered clinically important to patients with osteoarthritis, fibromyalgia, chronic low back pain, diabetic neuropathy, or postherpetic neuralgia.[47] When the present study accepted the threshold of a 10% to 20% improvement for minimum clinical importance, PCS was a clinically significant predictor of follow-up CPI in the subset of participants reporting moderate to severe baseline TMD pain intensity (ie, CPI > 68.7).

After effect modification by baseline CPI was taken into consideration, PCS score could contribute to a multifactorial model for predicting long-term pain, along with other factors not considered in this study's focused analyses. Such a comprehensive model could then be used to inform treatment and secondary prevention of long-term TMD pain.

Power and Sample Size

As a secondary data analysis, the present study had a fixed sample of 258 cases determined by applying the inclusion and exclusion criteria to the existing cohort of the Validation[38] and TMJ Impact Projects.[39] With statistically significant results for PCS score, the statistical power was necessarily adequate. For JFLS and MCS scores, which did not have statistically significant associations with follow-up CPI, the respective CIs did not include clinically significant changes (at least 10% change from mean baseline CPI, or 5.0 points).

Factors That May Have Reduced Observed Clinical Significance of Predictors

The long interval between baseline and follow-up data may partially account for

the limited clinical significance of PCS score and the absence of statistical significance of MCS and JFLS scores as predictors of long-term TMD pain intensity. The validated 6-month reference[43,48] for the CPI questions was used because this was the period used in the original GCPS questionnaire[5] and in the Validation Project.[38] Also, considering the fluctuating nature of TMD pain, it is possible that pain status at follow-up may differ from other times between assessments. However, a 6-month reference period, compared to shorter intervals, is thought to provide a better estimate of pain in the long interval between baseline and follow-up.

Potential Bias Associated With Partial Follow-up

Of the 513 eligible participants with baseline TMD pain in the Validation Project, 273 were followed up for the TMJ Impact Project and 240 were not followed up. Another 15 participants were excluded in the present study due to inconsistent or missing data, yielding the sample of 258 participants. Comparing included participants (n = 258) vs not included (n = 255) participants revealed a statistically significant 2.6-year difference in average age that was unlikely to have caused a clinically significant difference for the risk of long-term TMD pain, and the 0.39-point difference in JFLS score represented less than 5% of the range of baseline JFLS scores. Finally, the unchanged statistical significance of the study predictors in the sensitivity analysis gave the most compelling evidence that selective follow-up most likely did not introduce noteworthy bias.

Generalizability of Study Results

Participants in the Validation Project were enrolled according to the Standards for Reporting Diagnostic Accuracy (STARD)[49] guidelines, which have recommended using a sample of individuals with the target condition who are free of relevant comorbidities when first validating diagnostic criteria. Thus, it is possible that results could have differed in samples with other medical and pain comorbidities.

The Validation Project was designed to generalize its findings to TMD cases with the most common pain-related TMD diagnoses: TMJ soft tissue disorders (disc displacement) and TMJ hard tissue disorders (degenerative joint disease). With a large number of subjects recruited across the USA (New York, Minnesota, and Washington), the generalizability of the Validation Project results has been demonstrated.[41] No significant interactions ($P > .1$) were found between the three study predictors and three selected population characteristics: sex, age as a continuous measure, and age when the population was split into two categories. The authors conclude that these study results are generalizable to individuals having one or more pain-related TMD diagnoses, whether male or female, for ages in the range of this study.

Conclusions

In participants with moderate to severe baseline TMD pain intensity, higher baseline physical HRQoL (but not baseline mental HRQoL or baseline jaw functioning) was a statistically and clinically significant predictor of lower TMD pain intensity at 8 years follow-up. This secondary data analysis suggests PCS could contribute to a future multifactorial prediction model for long-term TMD pain for the purpose of improving clinical management and secondary prevention of ongoing TMD pain. Further research is needed to validate these findings.

Acknowledgments

Research reported in this publication was supported by the NIH/NIDCR: U01-DE013331, U01-DE019784, R90-DE023059, and the National Center for Advancing Translational Sciences of the National Institutes of Health Award Number UL1TR000114. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank Dr Donald R. Nixdorf and Dr Estephan J. Moana-Filho for project guidance and critical review. The authors report no conflicts of interest related to this study.

References

1. Okeson JP, de Kanter RJ. Temporomandibular disorders in the medical practice. J Fam Pract 1996;43:347–356.
2. LeResche L. Epidemiology of temporomandibular disorders: Implications for the investigation of etiologic factors. Crit Rev Oral Biol Med 1997;8:291–305.
3. National Institute of Dental and Craniofacial Research. Facial pain. http://www.nidcr.nih.gov/DataStatistics/FindDataByTopic/FacialPain/. Accessed 19 October 2017.
4. NIH Medline Plus. Chronic pain: Symptoms, diagnosis, and treatment. https://www.nlm.nih.gov/medlineplus/magazine/issues/spring11/articles/spring11pg5-6.html. Accessed 19 October 2017.
5. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain 1992;50:133–149.
6. Quartana PJ, Buenaver LF, Edwards RR, Klick B, Haythornthwaite JA, Smith MT. Pain catastrophizing and salivary cortisol responses to laboratory pain testing in temporomandibular disorder and healthy participants. J Pain 2010;11:186–194.
7. Johnson J, Johansson M, Rydén A, Houltz E, Finizia C. The impact of trismus on health-related quality of life and mental health. Head Neck 2015;37:1672–1679.
8. Dahlström L, Carlsson GE. Temporomandibular disorders and oral health-related quality of life. A systematic review. Acta Odontol Scand 2010;68:80–85.

9. De Leeuw R, Bertoli E, Schmidt JE, Carlson CR. Prevalence of traumatic stressors in patients with temporomandibular disorders. J Oral Maxillofac Surg 2005;63:42–50.
10. Epker J, Gatchel RJ. Coping profile differences in the biopsychosocial functioning of patients with temporomandibular disorder. Psychosom Med 2000;62:69–75.
11. Zakrzewska JM. Multi-dimensionality of chronic pain of the oral cavity and face. J Headache Pain 2013;14:37.
12. Velly AM, Look JO, Carlson C, et al. The effect of catastrophizing and depression on chronic pain—A prospective cohort study of temporomandibular muscle and joint pain disorders. Pain 2011;152:2377–2383.
13. McBeth J, Macfarlane GJ, Benjamin S, Silman AJ. Features of somatization predict the onset of chronic widespread pain: Results of a large population-based study. Arthritis Rheum 2001;44:940–946.
14. Lacey RJ, Lewis M, Sim J. Piecework, musculoskeletal pain and the impact of workplace psychosocial factors. Occup Med (Lond) 2007;57:430–437.
15. Linton SJ. A review of psychological risk factors in back and neck pain. Spine (Phila Pa 1976) 2000;25:1148–1156.
16. Melloh M, Elfering A, Egli Presland C, et al. Identification of prognostic factors for chronicity in patients with low back pain: A review of screening instruments. Int Orthop 2009;33:301–313.
17. Walton DM, Carroll LJ, Kasch H, et al. An overview of systematic reviews on prognostic factors in neck pain: Results from the International Collaboration on Neck Pain (ICON) project. Open Orthop J 2013;7:494–505.
18. Zakrzewska JM. Differential diagnosis of facial pain and guidelines for management. Br J Anaesth 2013;111:95–104.
19. Sanders AE, Slade GD, Bair E, et al. General health status and incidence of first-onset temporomandibular disorder: The OPPERA prospective cohort study. J Pain 2013; 14(suppl):T51–T62.
20. Fillingim RB, Ohrbach R, Greenspan JD, et al. Psychological factors associated with development of TMD: The OPPERA prospective cohort study. J Pain 2013;14(suppl):T75–T90.
21. Fillingim RB, Ohrbach R, Greenspan JD, et al. Potential psychosocial risk factors for chronic TMD: Descriptive data and empirically identified domains from the OPPERA case-control study. J Pain 2011;12(suppl):T46–T60.
22. Aggarwal VR, Macfarlane GJ, Farragher TM, McBeth J. Risk factors for onset of chronic oro-facial pain—Results of the North Cheshire oro-facial pain prospective population study. Pain 2010;149:354–359.
23. Lipton RB, Liberman JN, Kolodner KB, Bigal ME, Dowson A, Stewart WF. Migraine headache disability and health-related quality-of-life: A population-based case-control study from england. Cephalalgia 2003;23:441–450.
24. Picavet HS, Hoeymans N. Health related quality of life in multiple musculoskeletal diseases: SF-36 and EQ-5D in the DMC3 study. Ann Rheum Dis 2004;63:723–729.
25. Salaffi F, De Angelis R, Stancati A, Grassi W, Marche Pain, Prevalence Investigation Group (MAPPING) study. Health-related quality of life in multiple musculoskeletal conditions: A cross-sectional population based epidemiological study. II. The MAPPING study. Clin Exp Rheumatol 2005;23:829–839.
26. Tjakkes GH, Reinders JJ, Tenvergert EM, Stegenga B. TMD pain: The effect on health related quality of life and the influence of pain duration. Health Qual Life Outcomes 2010;8:46.
27. Barros Vde M, Seraidarian PI, Côrtes MI, de Paula LV. The impact of orofacial pain on the quality of life of patients with temporomandibular disorder. J Orofac Pain 2009;23:28–37.
28. Optum. SF-12v2 Health Survey. https://www.optum.com/optum-outcomes/what-we-do/health-surveys/sf-12v2-health-survey.html. Accessed 24 October 2017.
29. Ware J Jr, Kosinski M, Keller SD. A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220–233.
30. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–483.
31. Lipton RB, Hamelsky SW, Kolodner KB, Steiner TJ, Stewart WF. Migraine, quality of life, and depression: A population-based case-control study. Neurology 2000;55:629–635.
32. Mas AJ, Carmona L, Valverde M, Ribas B, EPISER Study Group. Prevalence and impact of fibromyalgia on function and quality of life in individuals from the general population: Results from a nationwide study in Spain. Clin Exp Rheumatol 2008;26:519–526.
33. Cheak-Zamora NC, Wyrwich KW, McBride TD. Reliability and validity of the SF-12v2 in the medical expenditure panel survey. Qual Life Res 2009;18:727–735.
34. Bair E, Ohrbach R, Fillingim RB, et al. Multivariable modeling of phenotypic risk factors for first-onset TMD: The OPPERA prospective cohort study. J Pain 2013;14(suppl):T102–T115.
35. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113:9–19.
36. Ohrbach R, Larsson P, List T. The jaw functional limitation scale: Development, reliability, and validity of 8-item and 20-item versions. J Orofac Pain 2008;22:219–230.
37. Ohrbach R, Fillingim RB, Mulkey F, et al. Clinical findings and pain symptoms as potential risk factors for chronic TMD: Descriptive data and empirically identified domains from the OPPERA case-control study. J Pain 2011;12(suppl):T27–T45.
38. Schiffman EL, Truelove EL, Ohrbach R, et al The Research Diagnostic Criteria for Temporomandibular Disorders. I: Overview and methodology for assessment of validity. J Orofac Pain 2010;24:7–24.
39. Schiffman EL, Ahmad M, Hollender L, et al. Longitudinal stability of common TMJ structural disorders. J Dent Res 2017; 96:270–276.
40. Schiffman E, Ohrbach R, Truelove E, et al. Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for Clinical and Research Applications: Recommendations of the International RDC/TMD Consortium Network and Orofacial Pain Special Interest Group. J Oral Facial Pain Headache 2014;28:6–27.
41. Schiffman EL, Ohrbach R, Truelove EL, et al. The Research Diagnostic Criteria for Temporomandibular Disorders. V: Methods used to establish and validate revised Axis I diagnostic algorithms. J Orofac Pain 2010;24:63–78.
42. Ohrbach R, Granger C, List T, Dworkin S. Preliminary development and validation of the jaw functional limitation scale. Community Dent Oral Epidemiol 2008;36:228–236.
43. Ohrbach R, Turner JA, Sherman JJ, et al. The Research Diagnostic Criteria for Temporomandibular Disorders. IV: Evaluation of psychometric properties of the Axis II measures. J Orofac Pain 2010;24:48–62.
44. EuroQol Group. EuroQol—A new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.
45. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. Stat Methods Med Res 2013; 22:278–295.
46. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008;9:105–121.
47. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001;94:149–158.
48. Von Korff M. Assessment of chronic pain in epidemiological and health services research: Epirical bases and new directions. In: Turk DC, Melzack R (eds). Handbook of Pain Assessment, ed 3. New York: Guilford, 2011:455–471.
49. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. Ann Intern Med 2003;138:W1–W12.