Patient Reported Outcome Measures (PROMs) used in temporomandibular disorders. A review of the literature.

Dina Taimeh[1,2], Richeal Ni Riordain[1,3], Stefano Fedele [1,4,5], Rachel Leeson [1]

1. Department of Maxillofacial Medicine and Surgery, UCL Eastman Dental Institute, University College London. London, UK
2. Department of Oral and Maxillofacial Surgery, Oral Medicine and Periodontology, Faculty of Dentistry, University of Jordan. Amman, Jordan
3. Oral Medicine Unit, Cork University Dental School and Hospital, Ireland.
4. Oral Medicine Unit, Eastman Dental Hospital, University College London Hospitals Trust. London, UK
5. NIHR University College London Hospitals Biomedical Research Centre. London, UK

Abstract

Aims: To identify the range of PROMs used in TMD studies, to summarise the available evidence of psychometric properties and to provide guidance for the selection of such measures. **Materials and methods**: A comprehensive search was conducted to retrieve the published articles in the time period between 2009-2018 containing a patient reported measure of the effects of TMD. Three databases were searched: Medline, Embase, and Web of Science. **Results**: 517 articles containing at least one PROM were included in the review. 58 additional studies were also located describing the psychometric properties of some tools in a TMD population. A total of 106 PROMs were identified and fell into the following categories: PROMs describing the severity of symptoms, PROM describing the psychological status, and PROMs describing the quality of life and general health. The most commonly used PROM was the Visual Analogue Scale. However, a wide range of verbal descriptors were employed. The Oral Health Impact Profile-14 (OHIP-14) and Beck Depression Inventory were the most commonly used PROMs describing the effect of TMD on quality of life and psychological status, respectively. Additionally, the OHIP (various versions) and The Research Diagnostic Criteria Axis ll questionnaires were the

instruments most repeatedly tested in a TMD population and undergone cross-cultural validation into several languages. **Conclusion**: a wide range PROMs have been used to describe the impact of TMD on patients. Such variability may limit the ability of the researchers and clinicians to evaluate the efficacy of different treatments and make meaningful comparisons.

Keywords: patient reported outcome measures, temporomandibular disorders, psychometric properties, review, quality of life.

Introduction

The term Temporomandibular Disorder (TMD) is a collective term embracing a number of clinical conditions that involve the masticatory musculature, the temporomandibular joint and associated structures [1]. It represents the most common cause for chronic pain in the facial region [2]. The associated symptoms include pain, restricted mouth opening, deviation in mandibular movements, clicking noises of the joint, headache in the temporal region, and psychological effects. [3,4]. Among the various categories which are grouped under the umbrella term, muscle problems represent the largest category [5]. Chronic pain may have severe distressing social and emotional effects. Indeed, depression, anxiety and negative beliefs about pain are not only linked to developing chronic pain, but also seem to contribute to worse outcomes from it [6]. Some initiatives such as The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) [7] and the Diagnostic Criteria for TMD (DC/TMD) [8] have therefore recommended the assessment of not only physical functioning, but also the psychological and emotional functioning associated with chronic pain.

Traditionally, healthcare has been assessed in terms of the technical and physiological outcomes of treatment [9]. In more recent times however, healthcare organisations are

striving to achieve services that are not only clinically effective and evidence-based, but also beneficial and effective to patients as judged from their own perspective [9,10]. In an attempt to increase efficiency without decreasing the humanity of the patient encounter, the use of Patient Reported Outcome Measures (PROMs) was proposed. These questionnaires could be a very powerful tool to bridge the need for gathering information in an efficient manner, complement the clinical decision making, and enhance communication between patients and physicians [11]. Many instruments exist that measure the intensity of pain, quality of life, psychological distress and disability. Some are generic, which are used in a wide range of conditions and settings, and some are condition-specific, which are designed specifically for the use of certain populations [12]. The aim of the current review was to identify the range of PROMs used in clinical studies of TMD patients and to review which PROMs have undergone psychometric testing in a TMD population in order to provide guidance for the selection of such measures.

Materials and methods

The search strategy

A comprehensive search was conducted in January of 2019 to retrieve the published articles that are concerned with the patient reported assessment of the effects of temporomandibular disorders. The articles were retrieved from three databases: Medline, Embase, and Web of Science. The employed search strategy consisted of the following MeSH terms and keywords: patient-reported outcome, outcome assessment, patient reported outcome measures, treatment outcome, patient centred outcome, patient defined outcome, subjective outcomes AND Temporomandibular joint disorder, Myofascial pain, temporomandibular joint dysfunction syndrome, TMD,

TMJD, facial myalgia, facial arthralgia, temporomandibular joint derangement, temporomandibular disc displacement.

Due to the vast number of articles retrieved, the search was restricted to the following 10-year time period: 2009-2018. Articles assessing the psychometric properties of the PROMs were located using the same search strategy. However, no time restrictions were applied (i.e., all the articles yielded from the search strategy were screened up to January 2019).

The included studies were clinical trials and observational studies of TMD (cross sectional and longitudinal) containing at least one PROM, articles reporting on the development or psychometric testing of a PROM in a TMD population and articles published in peer reviewed journals in the English language. The exclusion criteria eliminated studies containing clinical or radiological outcomes only, studies containing PROMs that report on the side effects after a specific intervention (e.g., complications of surgery), systematic and literature reviews, case reports, book chapters, conference proceedings, commentary or author opinion, animal studies, and studies with full text unavailable.

Data extraction

A study-specific Excel spreadsheet was used to aid with consistent data extraction. The following information was extracted: Study design, type of intervention (if any), number of participants, age range (or mean age), type of TMD, classification system used, the PROM used, the follow up time point (if any). Additional data were also extracted from studies that assess the psychometric properties of the PROMs in a TMD population, such as measures of validity, reliability, interpretability, and responsiveness.

Results

The initial search of the three mentioned databases yielded 3452 articles in total. After applying the exclusion criteria, 517 articles containing at least one PROM remained. Additional 58 articles were also found describing some form of psychometric testing, including cross-cultural validation. Most of the included studies employed a TMD classification system (64%, n=331), with the most commonly used system being the RDC/TMD criteria (50.68%, n=262), followed by the Wilkes classification system (7.9%, n=41) and the American Academy of Orofacial Pain Criteria (1.7%, n=9).

Patient Reported Outcome Measures commonly used in TMD

A total of 106 PROMs were identified after examining the included studies. The PROMs fell into three categories: PROMs describing the severity and improvement of symptoms, PROMs describing the psychological status and satisfaction, and PROMs describing the quality of life and general health. See table 1 for the identified PROMs and the frequency of use.

The most commonly used PROM was the Visual Analogue Scale with 59.5% of the trials using this instrument. However, various verbal descriptors were employed, such as: pain intensity, subjective chewing efficiency and quality of life. The rating scale associated with the VAS also varied, with most trials reporting results on a 100mm scale (or a 10cm scale). In a few studies, however, VAS was associated with scales ranging from 0-3, 0-4, 0-5, 0-6, and -5-5 scales (highlighting the possible misuse of the VAS).

Likert Point Scales and Numeric Rating Scales were also relatively commonly used (19.9% and 12.4% respectively). Similar to the VAS, the word descriptors varied for these PROMs, as did the length of the scales. The Point Scale mostly ranged from 3 to 7 points; however, 5 studies used an 11-point scale. For the NRS, the increments

of the scales included 0-10, 0-3, 0-5, 0-6 and 1-4. Among the other common PROMs were the Symptom Checklist-90 (SCL-90), and the Graded Chronic Pain Scale (GCPS) (8.7% and 8.1%, respectively).

Most of the other PROMs described the characteristics of pain, and the functional limitations incurred. A few PROMs described other symptoms associated with TMD, such as the Neck Disability Index (n=6) Tinnitus Handicap Inventory (n=2), Headache Impact Test-6 (n=2) and Food Intake Ability (FIA) index (n=1).

As for the PROMs assessing quality of life, Oral Health Impact Profile-14 was most frequently employed (5.61%). Most of the PROMs used to describe quality of life were generic instruments, except TMJ-Surgical-Quality of Life (TMJ-S-QoL) which is specific to TMD.

In total, 36 PROMs which described the psychological status of the participants were identified. The most frequently used PROM describing psychological distress was the Beck Depression Inventory (2.13%), followed by the Pain Catastrophizing Scale (1.93%).

Psychometric properties of PROMs used in TMD

Several PROMs identified in our search have some evidence of psychometric testing in a TMD population. The PROMs identified and their relevant psychometric evidence are detailed in table 2. The Research Diagnostic Criteria (RDC) Axis II tools and the Oral Health Impact Profile (OHIP) were the instruments most repeatedly tested in a TMD population and undergone cross-cultural validation into several languages. The search also identified a TMD-specific variant of OHIP; OHIP-TMD. The reported psychometric properties were internal consistency (Cronbach's $\alpha = 0.94$), test-retest reliability (Intraclass correlation coefficient= 0.805), convergent validity, content validity, known groups validity and responsiveness to change. One other variant was

also suggested for orofacial pain, where the authors omitted 10 items from the original tool and added two items relevant to facial pain patients (Cronbach's α= 0.97).

Discussion

The recent growth of the adoption of PROMs into healthcare settings reflects the emphasis placed by health institutes on the importance and relevance of the patient perspective in improving the quality of healthcare. They are a shift from the more traditional indicators of treatment success, such as mortality rate, post-surgical infection rates and readmissions [13]. Additionally, using PROMs improves communication between the clinician and the patient, which in turn may improve satisfaction, and adherence to treatment [14]. Although PROMs are now commonly incorporated in the scientific literature, as outcomes in clinical trials concerning TMD for example, a uniform set of outcomes or instruments is not routinely used. This limits our ability to compare outcomes of these clinical trials across the various studies conducted.

Kavchak et al provided an assessment of the psychometric properties of some tools in a TMD population in 2014. The group was able to identify 13 papers describing some form of psychometric analysis for 8 tools. They reported in their review that few PROMs reported for use in TMD patients have undergone rigorous analysis and with complete psychometric properties established [15]. Aguiar et al, also examined the psychometric properties of 10 common condition-specific PROMs and had similar conclusions where they note the need for further studies on psychometric properties [16].

In the present review of 517 studies, 106 PROMs were identified which were used to assess the effects of TMD on patients, and additional 58 papers that tested the psychometric properties of some tools in a TMD population including cross-cultural

adaptation. The most used PROM was the Visual Analogue Scale. The pain VAS mimics the continuous visual analogue scales developed to measure well-being in the psychology domain [17]. It is relatively acceptable to patients [18], and widely used in diverse adult populations [19]. Other reviews of the literature have also reported that VAS is the most widely used PROM in Oral Medicine populations [20] such as Oral Lichen Planus [21] and Burning Mouth Syndrome [22]. This widespread use can be rationalised in light of the relative ease of administration, low administrative burden required and acceptability to patients [23]. The wide variety of word descriptors associated with it, however, could result in heterogeneity of the results, and difficulty of data pooling. The VAS may also have been misnamed in a small number of studies, where different increments were utilised (for example 0-3). VAS is usually displayed as 10-cm line that represents a continuum between the two ends of the scale [24]. Hence, the scores on such scales may be better labelled as likert-point scales.

The most frequently used oral-health quality of life PROM in our review was the Oral Health Impact Profile-14. The items for OHIP were generated following interviews with patients from private dental practice, primary care clinics and prosthetic clinics in a dental hospital [25], therefore, it may not be specific enough for patients with TMD to detect the impact of the condition on their daily lives. The TMD variant (OHIP-TMD) has good internal consistency reliability and test-retest reliability according to the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) criteria [26]. First proposed in 2011 [27] with further validation presented in 2015 [28], this measure is still relatively new compared to OHIP-14, which might explain the popularity of the latter in TMD research so far.

The current search has highlighted the scarcity of TMD-specific quality of life and psychological status PROMs. Several have been created to describe the symptoms

of TMD, such as Jaw Functional Limitation Scale (JFLS), Mandibular Function Impairment Questionnaire (MFIQ), Jaw Disability Checklist (JDC), and the Jaw Pain and Function (JPF)-Questionnaire. However, PROMs describing other dimensions of the condition are still lacking and most clinical trials have used generic PROMs to describe the quality of life and psychological status of the patients. Condition-specific PROMs are more sensitive and with greater discriminatory ability to detect small changes over time [29,30].

Numerous PROMs are used to describe the impact of TMD on patients. Such variability may limit the ability of researchers and clinicians to evaluate the efficacy of different treatments, data pooling and making meaningful comparisons. The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) has tried to address this issue in chronic pain trials [7]. This initiative recommends evaluating the following aspects: pain intensity, physical functioning, emotional functioning, participant ratings of global improvement and satisfaction with treatment, symptoms and adverse events, and participant disposition. It also recommends the use of certain PROMs to unify the results among clinical trials. The Research Diagnostic Criteria has also been proposed to provide a comprehensive diagnostic and classification system for the subtypes of temporomandibular disorders. The criteria, first proposed in 1992 [31], was updated in 2014 following a series of workshops to include an expanded taxonomic classification structure to include common and less common TMDs. Additionally, its second axis was expanded by adding new instruments to evaluate pain behaviour, psychological status, and psychosocial functioning [8]. Our results have highlighted that this classification system is popular among researchers, as 50.68% of the studies confirmed the diagnosis of TMD based on axis I of these criteria. However, fewer studies used the complete list of PROMs recommended in axis II. The length of

the proposed questionnaires may discourage some researchers. Additionally, the primary objective of a trial might involve other clinical or radiological outcomes, therefore, a comprehensive evaluation of psychosocial functioning may not be crucial to the researchers. It is, however, also important to mention that a core outcome set for clinical trials in TMD is currently under development [32]

It was also noted that while some studies employed subjective measures to assess function, other studies, or indeed the same study employed physical measures as well. For example, the VAS and NRS were used to describe not only pain intensity, but also subjective restrictions to function, such as limitation to mouth opening, difficulty in chewing and diet restrictions. In other instances, objective, operator-measured physical outcomes were used, such as maximum mouth opening and mandibular range of motion. Both physical and patients reported measures are essential in the assessment of temporomandibular disorders. Hence TMD classification criteria, for example DC/TMD, utilise the outcome of the clinical exam in addition to a patient completed symptom questionnaire to establish a diagnosis. Additionally, Loh et al report in their systematic review of trismus instruments, that the correlation between subjective and objective measures was overall strong, and the findings of some studies which used objective measures, were in line with studies measuring trismus subjectively. [33]

The present review was limited to studies in the English language found in the three mentioned databases. Indeed, the results of the search might be different should studies in other languages be included, or the search expanded to other databases with no time restrictions.

A comprehensive search was carried to locate the papers which test the psychometric properties of the different tools in a TMD population. This paper, however, did not

conduct a formal assessment of the included studies, where the methodology and adequacy of these properties were assessed. Nonetheless, a detailed summary of these papers was presented in table 2 to enable the readers the judge the suitability of each PROM to their own setting.

Conclusions

Condition specific PROMs to assess the psychological status and quality of life of TMD patients are needed. The scarcity of such measures is reflected by the popularity of generic PROMs used in TMD research. While these may be useful in comparisons between different populations, they may lack the sensitivity and discriminatory ability in specific conditions. The use of a collection of concise and psychometrically sound measures may also promote consensus in TMD literature and provide a more robust basis for comparisons and data pooling.

Key findings

- The VAS, OHIP-14, and BDI were the most commonly used PROMs to describe the pain intensity, oral health-related quality of life and psychological status of the participants in TMD studies, respectively.
- A wide variety of PROMs are used in TMD research, potentially influencing the ability to pool data and make meaningful comparisons of different treatment modalities.

Acknowledgements

Author contributions: DT: design, data extraction, data analysis and paper writing. RNR: project supervision, design and critical revision of manuscript. SF: project supervision, design and critical revision of manuscript. RL: project supervision design and critical revision of manuscript.

The authors would like to acknowledge the financial support provided by the University of Jordan, in the form of a PhD student sponsorship awarded to DT.

The authors declare no conflict of interest regarding this piece of research.

Table 1. PROMs identified and their frequency of use.

Name of PROM	Frequency of use
Severity of symptoms and improvement	
Visual Analogue Scale (VAS)	308
Point Scales	103
Numeric Rating Scale (NRS)	64
Symptom Checklist-90 (SCL-90)	45
Graded Chronic Pain Scale (GCPS)	42
Jaw Functional Limitation Scale (JFLS)	16
Mandibular Function Impairment Questionnaire (MFIQ)	16
McGill Pain Questionnaire	13
Fonseca anamnestic index (FAI)	12
The West Haven-Yale Multidimensional Pain Inventory (WHYMPI)	12
Adjectival scale	12
RDC/TMD Axis II	11
Verbal Rating Scale	10
Helkimo anamnestic dysfunction index	10
Jaw Disability Checklist (JDC)	9
Symptom Severity Index (SSI)	8
Brief Pain Inventory (BPI)	7
Neck Disability Index	6
Characteristic Pain Intensity (CPI)	4
3Q/TMD	3
Brief Symptom Inventory (BSI)	3
The Pain Related Self-Statement Scale	3
Chronic pain grade	2
Coloured Analogue Scale (CAS)	2
Headache Impact Test-6	2
Limitations in Daily Functions-Temporomandibular Disorders Questionnaire (LDF-TMDQ)	2
Manchester Orofacial Pain Disability Scale (MOPDS)	2
Pain Stages of Change Questionnaire (PSOCQ)	2
ProTMDMulti	2
The Oral Behavior Checklist (OBC)	2
Tinnitus Handicap Inventory	2

Widespread Pain Index (WPI)	2
Craniofacial Pain and Disability Inventory	1
Food Intake Ability (FIA) index	1
Mann assessment of swallowing ability (MASA) score	1
PRISM (Pictorial Representation of Illness and Self-Measure)	1
Screening for Somatoform Symptoms (SOMS-7)	1
Symptom Interference Questionnaire – Revised (SIQR)	1
The Battery for Health Improvement	1
The Gracely Pain Scale	1
the Jaw Pain and Function (JPF)-Questionnaire	1
The Pain Behavior Questionnaire	1
The Patient Specific Functional Scale (PSFS)	1
The Universal Pain Assessment Tool (UPAT)	1
Visual Faces Pain Scale (FPS)	1
Zerssen complaint list	1
Quality of Life, general health and effect on daily life questionnaires	
Oral Health Impact Profile-14 (OHIP-14)	29
Pittsburgh Sleep Quality Index	17
The Short Form 36 Health Survey (SF-36)	14
Oral Health Impact Profile (OHIP)	9
General Health Questionnaire-7 (GHQ-7)	7
The Short Form 12 Health Survey (SF-12)	5
Epworth Sleepiness Scale	4
TMJ-Surgical-Quality of Life (TMJ-S-QoL)	4
WHO QoL- brief	4
EQ-5Dm	3
Health Assessment Questionnaire	2
OHQoL-UK	2
Oral Health Impact Profile-OFP (OHIP-OFP)	2
RAND-36 health survey	2
University of Washington QOL (UW-QOL)	2
Youth Self report	2
Child Perception's Questionnaire	1
General Oral Health Assessment Index (GOHAI)	1
Insomnia Severity Index (ISI)	1
Michigan Oral Health-related Quality of Life Scale – Child Version (MOHRQoL-C)	1
The Flanagan Quality of Life Scale	1
The Life Experiences Survey (LES)	1
The Multidimensional Fatigue Inventory-Short Form	1
The Sleep Assessment Questionnaire (SAQ)	1
Psychological Status and Satisfaction	
Beck Depression Inventory (BDI)	11
Pain Catastrophizing Scale (PCS)	10

Hospital Anxiety and Depression Scale (HADS)	8
State-Trait Anxiety Index (STAI)	8
Patient Health Questionnaire-9 (PHQ-9)	7
Generalised Anxiety Disorder-7 (GAD-7)	6
Centre for Epidemiological Studies Depression scale-20 (CESD)	5
Coping Strategies Questionnaire (CSQ)	5
PTSD Check List–Civilian (PCL-C)	5
The Perceived Stress Scale (PSS)	4
Beck Anxiety Inventory (BAI)	3
Chronic Pain Self-Efficacy Scale (CPSS)	2
Eysenck Personality Questionnaire-Revised (Short Form, EPQ-R)	2
Life Orientation Test 12-Revised	2
Lipp's Stress Symptoms Inventory for Adults (LSSI)	2
Sense of coherence-29 (SOC-29)	2
The Pennebaker Inventory of Limbic Languidness: The Kohn Reactivity Scale	2
Behavioral Rating Scale (BRS)	1
Columbia Classification Algorithm of Suicide Assessment (C-CASA)	1
Coping Pain Questionnaire (CAD)	1
Dental Anxiety Scale (DAS)	1
Depression, Anxiety and Stress Scales-21 (DASS-21)	1
Fear Avoidance Belief Questionnaire (FABQ)	1
Illness Perception Questionnaire –Revised (IPQ-R)	1
Irrational Attitudes Questionnaire	1
Miller Behavioral Style Scale [MBSS]	1
Millon Behavior Medicine Diagnostic survey	1
Minnesota Multiphasic Personality Inventory (MMPI)	1
NEO-Five Factor Inventory (NEO-FFI)	1
Pain Coping and Cognition List (PCCL)	1
Screening for Somatoform Symptoms (SOMS-7)	1
Survey of Pain Attitude (SOPA-35)	1
Tampa Scale for Kinesiophobia (TSK-11)	1
The Group Health Association of America (GHAA) Consumer Satisfaction Survey	1
The Profile of Mood States-Bipolar (POMS-Bi)	1
The Satisfaction With Life Scale	1

Table 2. Details of the psychometric properties of some PROMs in a TMD population.

PROM	Author and year	Domains/factors	Number of TMD patients	Psychometric testing
Central Sensitization Inventory (CSI)- Italian	[34]	-	37	Cross-cultural adaptation Structural validity: Exploratory factor analysis Construct validity: Pearson's correlation with: 11 point NRS for pain intensity =0.427, SF 36= −0.479, HADS=0.706, Pain Self-Efficacy Questionnaire= −0.618. All have significant correlations. Internal consistency: Cronbach's α= 0.87
Centrality of pain scale- Chinese	[35]	-	166	Cross cultural adaptation Internal consistency: Cronbach's α= 0.942 Test-retest (30 patients- 1week): ICC= 0.815 - 0.929. Construct validity: Exploratory factor analysis (EFA)- 1 factor Convergent validity: Pearson's correlation with: Catastrophizing Scale (r=0.57) and Pain Self-Efficacy Questionnaire (r= –0.42). Both have significant correlation.
Child perception questionnaires **CPQ 8-10 (years)** **CPQ 11-14 (years)**	[36]	- Oral symptoms - Functional limitations - Emotional well-being - Social well- being	547	Criterion validity: Spearman's correlation with pain scores (obtained from Question 3 of the RDC/TMD Axis II). CPQ 8-10: r= 0.18/ non sig, CPQ 11-14: r=0.32/sig. Discriminant construct validity Correlational construct validity Internal reliability (internal consistency)- CPQ 8-18: Cronbach's α= 0.93, CPQ 11-14: Cronbach's α= 0.94
Craniofacial Pain and Disability Inventory (CFPDI)- Spanish	[37]	- Pain and disability - Jaw functional status	192	Test-retest reliability (106 patients, 12 days): ICC= 0.90 Internal consistency: Cronbach's α= 0.88 Construct validity by exploratory factor analysis: 2 factors Responsiveness: SEM= 2.4

				Convergent validity: Pearson's correlation with VAS= 0.46, PCS (r=0.46), TSK-11 (r=0.40), NDI (r=0.65), HIT-6 (r= 0.38). All have significant association.
Craniofacial Pain and Disability Inventory- Brazilian Portuguese	[38]		100	Cross cultural adaptation Internal consistency: Cronbach's α= 0.77-0.86 Construct validity: Pearson's correlation with PCS (0.69), TSK-TMD (0.68), NDI (0.40), MFIQ (0.74), and pain-related disability (0.75). All have significant correlation Structural validity: Confirmatory Factor Analysis- 3 factors Test retest (60 patients- 1 week): ICC= 0.97
EQ-5D-5L	[39]	- Mobility - Self-care - Usual activities - Pain/ discomfort - Anxiety/ depression	66	Convergent validity- Spearman's Rho with MPI for each subscale.
Fonseca anamnestic index (FAI)	[40]	-	94	Structural validity: Exploratory factor analysis Overall correlation between items : Spearman's correlation- Some items showed good correlation, but not all items were correlated, suggesting more than one dimension in the FAI. Internal consistency: Cronbach's α= 0.7 Rasch analysis
	[41]		700 – normal population, assuming 40% are TMD	Confirmatory Factor Analysis (CFA) Convergent validity - The average variance extracted (AVE)= 0.513, The composite reliability (CR)= 0.878 Internal consistency: Cronbach's α = 0.745 Reproducibility (62 patients- 1 week): Kappa = 0.89 Concurrent validity- Correlation analysis with MFIQ: r = 0.66

				(Questions 8 and 10 were below the adequate values. Thus, these questions were excluded from the original model)
Jaw Disability Checklist (JDC) **Characteristic Pain Intensity (CPI)** **Symptom Checklist 90-Revised (SCL-90-R)** **Oral Health Profile-14 (OHIP-14)- Turkish** **Short Form 36 Item Health Survey (SF-36)- Turkish**	[42]	SCL-90: - Somatisation - Obsessive-compulsive Interpersonal sensibility - Depression - Anxiety - Anger-hostility - Phobic-anxiety - Paranoid ideation - Psychoticism - Other items	104	Internal consistency: Cronbach's α JDC= 0.76 CPI= 0.79 SCL-90-R- somatisation= 0.87 SCL-90-R- depression= 0.93 OHIP-14= 0.86 SF 36- physical health= 0.83 SF-36- mental health= 0.82
Jaw Function Limitation Scale-20,8 **(JFLS-20, 8)**	[43]	- Mastication - Vertical jaw mobility - Emotional and verbal expression	31	Fitness of model/ item reliability: Rasch methodology Temporal stability (1-2 weeks): concordance correlation coefficient- JFLS 20= 0.87, JFLS 8 = 0.81 Internal consistency: Cronbach's α- JFLS 20 = 0.95, JFLS 8= 0.87 Correlation of Subscales: JFLS 20= 0.9422
	[44]		219	Factor analysis Model fitness: Rasch methodology Construct validity: correlation with Jaw Symptom Index = 0.57, SCL-90= 0.02, GCPS Pain Interference = 0.26, GCPS Characteristic Pain Intensity = 0.49 CPI, and STAI= 0.17
Jaw pain and function (JPF)-German	[45]	- Jaw pain - Jaw function	137	Cross cultural adaptation Concurrent construct validity (97 patients) - Pearson's correlation with maximum inter-incisal distance, r= -0.213. Significant correlation. Test-retest reliability (40 patients- 1 day and 1 week) - Pearson's correlation, 1 day: r= 0.91, 1week: r=0.93. Both are significant.

				Internal consistency of verbal subscales- Cronbach's α: Pain score = 0.85, ADL score = 0.94, function score = 0.68.
Manchester Orofacial Pain Disability Scale	[46]	- Physical - Psychological	171	Internal consistency- Cronbach's α: Physical disability construct = 0.78, psychosocial disability construct= 0.92 Item correlation: values between 0.43 and 0.80 Construct validity Factor analysis: 2 factors
	[47]		50	Cross cultural validation Internal consistency: Cronbach's α = 0.9 Test retest (reproducibility)- 15-20 days: ICC = 0.924 Criterion validity- Correlations with OHIP-14, r= 0.857 and VAS for pain intensity, r= 0.758. Both are significant. Inter-item correlation Factor analysis: 2 factors
Mandibular Function Impairment Questionnaire (MFIQ)	[48]	- Functional capacity - Feeding	95	Convergent validity Internal consistency: Cronbach's α= 0.63 to 0.95. Factor analysis: 3 factors

MFIQ- Chinese	[49]		352	Cross-cultural adaptation Internal consistency: Cronbach's α for factor 1: 0.925, for factor 2= 0.72 Test retest (78 patients - 7days): ICC for factor 1= 0.895, for factor 2= 0.720 Content validity: evaluated by twenty dentists and five physical therapists. Construct validity: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) - 2 factors. Face validity: consensus between 8 specialists
MFIQ- Portuguese	[50]		249	Factorial validity- confirmatory factor analysis: 2 factors Internal consistency: Cronbach's α for functional capacity dimension= 0.874, for feeding dimension= 0.918. Intra-rater reproducibility (62 patients – 1 week): ICC for functional capacity dimension 1 = 0.895, for feeding dimension= 0.825. Temporal stability (test retest reliability): Pearson's correlation for dimension 1 r= 896, for dimension 2 r= 0.826. Face validity: evaluated by six dentistry professionals (specialists on temporomandibular disorders) and three experts of the English language. Content validity: assessed by 21 dentists with expertise in temporomandibular disorders. Convergent and discriminant validity were assessed, respectively, by the average variance extracted (AVE), composite reliability (CC) and bivariate correlations between factors.
Multidimensional Pain Inventory (MPI)- Spanish	[51]	- Pain impact - Responses by significant others - Activities	114	Cross-cultural adaptation Internal consistency: Cronbach's α > 0.7 for all items. Confirmatory Factor Analysis

MPI - Brazilian	[52]		31	Convergent validity: Average variance extracted and composite reliability Internal consistency: Cronbach's α= 0.80–0.94 Content validity ratio (CVR): 15 experts in the field of dentistry. Construct validity- confirmatory factor analysis
Oral Behaviour Checklist- Portuguese	[53]	- Activities during sleep - Activities during waking hours	120	Cultural adaptation Test retest (120 patients- 2 weeks)- ICC= 0.998 Temporal stability: weighted Kappa = >0.946 Item agreement between English and Portuguese OBC: weighted Kappa = >0.934 Internal consistency: Cronbach's α= 0.64 Convergent and discriminant validity
OBC- Dutch	[54]		155	Cross cultural validity Test-retest reliability (35 patients- 2weeks): ICC= 0.86 Concurrent validity: Spearman's correlation with Dutch Oral Parafunctions Questionnaire r = 0.757, RDC- CPI r= 0.069, Dutch SCL-90 Depression r= 0.485, somatisation r= 0.312, anxiety r= 0.448, Stress 7 item questionnaire r= 0.433. All have significant correlations except with RDC-CPI. Correlations between individual items: 0.389 to 0.892
Oral Health Impact Profile-49 (OHIP-49)- German	[55]	- Functional limitation - Physical pain - Psychological discomfort - Physical disability - Psychological disability - Social disability - Handicap	67	Cross cultural validation Groups validity: Point-biserial correlations Responsiveness (1 month): Effect size calculation by paired t-test.
OHIP- 5,14,21-German	[56]		175	Validity and internal consistency: Cronbach's α = 0.65–0.92 Responsiveness: standardized effect size = 0.55-0.95

				Construct validity: Point-biserial correlations
OHIP-49- Swedish	[57]		30	Test retest reliability: ICC= 0.87-0.98 Construct validity: Spearman's correlation with JFLS (r= 0.76), SCL-90 (r= 0.65), self-reported health (r= 0.61). Internal reliability: Cronbach's α = 0.83-0.91
OHIP- Italian	[58]		124	Cross-cultural validation Content validation: group of experts Internal consistency: Cronbach's α= 0.71-0.86 Construct validation: known-groups analysis Criterion-related validation Exploratory factor analysis: 7 factors
OHIP-5,14,48-Dutch	[59]		245	Internal consistency: Cronbach's α- OHIP-48= 0.96, OHIP-14= 0.9, OHIP-5= 0.67 Test retest reliability (64 patients- 2 weeks): ICC- OHIP-48= 0.82, OHIP-14= 0.8, OHIP-5= 0.69 Construct validity Convergent validity: Spearman's rho with Pain-related disability score- OHIP-48= 0.46, OHIP-14= 0.46, OHIP-5= 0.39, and Self-reported oral health status- OHIP-48= 0.28, OHIP-14= 0.19, OHIP-5=0.21 Group validity: T tests between patients with and without complaints, and Spearman's rho (with CPI and biting activities)
OHIP-30-OFP	[60]		121	Internal consistency: Cronbach's α= 0.97
OHIP-TMD	[27]		110	Convergent validity: Spearman's Rho correlation with MPI = 0.751, VAS = 0.576. Without the 2 new items. Both are significant. Internal consistency: Cronbach's α= 0.942. Without the 2 new items.
OHIP-TMD	[28]		76	Face and Content validity: Focus groups of patients and a panel of specialists. Content validity index= 0.64 for patients, 0.82 for professionals.

				Known groups validity: t-tests of the means between patients and controls. Responsiveness to change: Paired, two tailed, t-tests to calculate effect size (OHIP-TMDs versus OHIP-49) Test retest reliability: ICC= 0.805 Internal consistency: Cronbach's α = 0.95 at baseline, 0.96 at follow up.
OHIP- TMD-Chinese	[61]		156	Cross cultural validation Internal consistency Cronbach's α = 0.917 Test retest (30 patients - 2weeks) : ICC= 0.899 Structural validity: Factor analysis- 5 factors Convergent validity: Global rating of oral health question= 0.548. Significant correlation.
Pain Disability Index	[62]	- Family/Home responsibilities - Recreation - Social activity - Occupation - Sexual behaviour - Self-care - Life-support activities	197	Factor structure
Pain related limitations of daily functions (LDF-TMDQ)- Japanese	[63]	- Limitation in executing a certain task - Limitation of mouth opening - Limitation of sleeping	456	Factor validity- Exploratory factor analysis: 3 factors - Confirmatory factor analysis Convergent validity Discriminant validity: Spearman correlations with Pain VAS, Japanese dental version of McGill Pain Questionnaire, HADS, Eysenck Personality Questionnaire short form and Diet VAS. Internal consistency: Cronbach's α = 0.81, and split-half estimation (Guttmann method) r = 0.76, ($P < 0.05$)
Pain resilience scale- Chinese	[64]	- Cognitive/Affective Positivity	152	Cross cultural validity Confirmatory Factor Analysis: 2 factors.

		- Behavioural perseverance		Internal consistency: Cronbach's α = 0.92 Test retest (30 patients - 2weeks): ICC = 0.92 Convergent validity: Spearman's correlation with Connor-Davidson Resilience Scale = 0.61 to 0.65 and TSK-TMD= -0.46 to -0.41
Pain-Related Control Scale (PRCS)	[65]	- Helplessness - Resourcefulness	44	Internal consistency: Cronbach's α = 0.83, 0.77 Convergent validity Discriminant validity Factor analysis Stability (Test-retest): PRCS-Helplessness= 0.86, PRCS-Resourcefulness= 0.88
Pain-Related Self Statements Scale (PRSS)	[65]	- Catastrophizing - Coping	44	Internal consistency: Cronbach's α= 0.92, 0.88. Convergent Discriminant validity Factor analysis Stability (Test-retest): PRSS-Catastrophizing= 0.87, PRSS-Coping= 0.77
Pittsburgh Sleep Quality Index (PSQI)	[66]	- Subjective sleep quality - Sleep latency - Sleep duration - Habitual sleep efficiency - Sleep disturbances - Use of sleeping medication - Daytime dysfunction	609	Exploratory factor analysis: 1 factor Model fit: Confirmatory factor analysis Internal consistency: Cronbach's α= 0.75 Inter-item correlation: Pearson correlation coefficients = 0.3 Test-retest reliability: ICC= 0.86 Convergent validity: Spearman's rho coefficient with questions from the GHQ, Q1= 0.43, Q2= 0.48.
PRISM (pictorial representation of illness and self-measure)	[67]	-	70	Construct validity: Pearson's correlation with GCPS (disability subscale) = –0.60, GCPS (PI subscale) = –0.55, HADS-D= –0.21, HADS-A= –0.21, Insomnia Severity Index = –0.41. Significant correlation with GCPS subscales and the ISI. Nonsignificant correlations with HADS subscales

PRISM (German to Portuguese)	[68]		42	Cross cultural translation Content validity: Pearson correlations with Numerical Pain Scale (NPS) 0-10 (moderate – 0.42), Insomnia Severity Index (week –0.24), HADS-A (week –0.25), HADS-D(week –0.22), Temporal stability (30 patients- 3days): ICC= 0.991
ProTMDMulti	[69]	-	30	Criterion Validity: Spearman R with Helkimo Di = 0.65. Significant correlation Construct Validity: Comparison results between pre- and post-treatment and comparing the TMD group to the control group
RDC/TMD- Axis II	[70]	Graded chronic pain, depression, somatization with and without pain, jaw disability checklist	362	Concurrent validity of SCL-90- depression: Pearson correlations with BDI= 0.69, and Centre for Epidemiologic Studies for Depression= 0.78 Internal consistency reliability: Cronbach's α, SCL90= 0.91, Non- Specific Physical Symptoms= 0.82, CCPS= 0.71. Construct validity of the Non-Specific Physical Symptoms Scale: Exploratory factor analysis- 2 factors Clinical utility: sensitivity of 0.91 and specificity of 0.41.
RDC/TMD Axis II	[71]		626	Internal consistency: Cronbach's α SCL-90-Depression= 0.91, Nonspecific Physical Symptoms, with pain items= 0.84, GCPS-CPI=0.84, GCPS- Activity Interference 0.95. Convergent validity: SCL-90- Depression: Lin's correlation concordance coefficient (CCC) with The Centre for Epidemiologic Studies-Depression instrument (CESD) = 0.85, and with SF12 = −0.70. SCL-90 Non- Specific Physical symptoms: CCC with GHQ-28= 0.45 and CESD= 0.56. GCPS-CPI: CCC with MPI= 0.65. GCPS- Activity Interference: CCC with MPI= 0.52 Test-retest reliability (75 patients-2 weeks): SCL-90-Depession: CCC= 0.63- 0.78.

				SCL-90 Non- Specific Physical symptoms: CCC= 0.63 – 0.78. GCPS-CPI (3 days): CCC= 0.91. GCPS- Activity Interference: CCC = 0.89 GCPS-Chronic pain grade: weighted kappa = 0.87 Discriminant validity: Lin's correlation concordance coefficient with MPI. Criterion validity Clinical utility of the Depression instrument by calculating PPV, NPV
RDC/TMD Axis II- Portuguese	[72]		55	Internal consistency: Cronbach's α= 0.72 Reliability: Kappa, 0.73 to 0.9 Test retest (45 patients- 2 weeks): Cohen Kappa scale/ for axis 1. Spearman's rank correlation = 0.727-0.821. Concurring validation: Spearman's correlation with Oral Impacts on Daily Performances= 0.306-0.602, OHIP-14= 0.336- 0.598
RDC/TMD axis II - German	[73]		378	Cross cultural adaptation Test-retest reliability (27 patients- 1-2 weeks): ICC- Jaw Disability List (JDL)=0.76, GCPS= 0.92 Internal consistency: Cronbach's α- JDL= 0.72, GCPS = 0.88. Construct validity: Rank correlation with self-reported oral health, OHIP-G, self-report of oral habits, MPI
RDC/TMD axis II Malay	[74]		40	Cross cultural validity Internal consistency: Cronbach's α- GCPS= 0.77, Nonspecific Physical Symptoms= 0.71, Depression= 0.88. Test-retest reliability (40 patients - 1 week): ICC- GCPS = 0.97, Nonspecific Physical Symptoms= 0.94, Depression= 0.95. Discriminant validity: t test of means between patients with pain symptoms and symptoms free. SEM

Multimedia Version of the RDC/TMD Axis II- Portuguese	[75]		30	Internal consistency: Cronbach's α = 0.94 Convergent validity: Spearman's rank correlation Reproducibility (1 day): Spearman´s rank correlation test= 0.670-0.913.
Screening for Somatoform Symptoms (SOMS-7)	[76]	- Somatization symptom count - Somatization severity index	58	Internal consistency: Cronbach's α = 0.88
Self-medication questionnaire	[77]	- Symptoms - Opinion about self-medication.	110	Face validity (content validity): interviews with 10 patients and expert opinion. Internal reliability: Cronbach's α =0.844 Exploratory factor analysis: 2 factors Reproducibility (11 patients-15 days): weighted Kappa coefficient=0.81
Short Form 36 Item Health Survey (SF-36)	[78]	1) limitations in physical activities because of health problems; 2) limitations in social activities because of physical or emotional problems; 3) limitations in usual role activities because of physical health problems; 4) bodily pain; 5) general mental health, 6) limitations in usual role activities because of emotional problems; 7) vitality and 8) general health perceptions.	146	Correlation of the SF-36 versus the Axis II scales: Spearmen coefficient (r). All items and subscales are significantly correlated with the exception of the jaw disability checklist when crossed with the mental scales of SF-36.
Social support and Pain Questionnaire (SPQ)- Chinese	[79]	-	118	Translation and cross-cultural adaptation Internal consistency: Cronbach's α = 0.926 Test retest reliability (2 weeks): ICC= 0.784

				Construct validity: Exploratory factor analysis- 1 factor model Convergent validity: Spearman's rank correlation with Global oral health question = 0.624. Significant correlation.
Social Support Scale	[80]	-Perceptions of social support -Satisfaction with social support.	92	Internal consistency: Cronbach's α =0.39-0.73 Criterion validity Construct validity: Correlation with The Centre for Epidemiologic Studies Depression Scale (CESD), Profile of Mood States (POMS), The Taylor Manifest Anxiety Scale (TMAS)
Symptom severity index- modified (SSI)	[81]	- Jaw pain - Temple pain	108	Internal consistency: Cronbach's α = 0.96 Dimensionality- exploratory factor analysis: 2 factors Test-retest reliability (55 patients- 2–48 hours): ICC= 0.97 Between-item correlation: substantial but variable
Tampa scale for kinesiophobia (TSK-TMD)-from original Dutch to English	[82]	- Activity avoidance - Somatic focus	301	Cross cultural adaptation Factor structure- Confirmatory factor analysis: 2 factors Test-retest reliability (4 weeks-58) : ICC= 0.73 Convergent validity: Pearson Correlation with the Catastrophizing scale of the Coping Strategies Questionnaire (Dutch version) = 0.23 Internal consistency: Cronbach's α = 0.83
TSK-TMD- Chinese	[83]		160	Translation and cross-cultural adaptation Internal consistency: Cronbach's α =0.919 Test-retest reliability (30 patients- 2 weeks): ICC= 0.797 Content validity: Interviews with patients and an expert panel Construct validity: exploratory factor analysis (EFA)- 2 factors

				Convergent validity: Pearson Correlation with Global oral health question =0.458–0.563
TSK-TMD-Brazilian Portuguese	[84]		100	Cross cultural validity Internal consistency: Cronbach's α = 0.78 Test retest: ICC= 0.51-0.75 Structural validity: confirmatory factor analysis- 2 factors Construct validity: Spearman's rank correlation with PCS= 0.48, PHQ-8= 0.38, MFIQ= 0.43 Convergent validity/ Discriminant validity: Average variance extracted
TMD-Pain Screening Instrument **Long Version (LV)** **Shot Version (SV)**	[85]	-	504	Internal reliability: Cronbach's α, LV= 0.93, SV= 0.87 Rasch analysis Sensitivity =99 % and Specificity =97% Exploratory-factor analysis (EFA) Temporal stability: ICC- LV= 0.79, SV= 0.83.
VAS score of the PSA (Patient specific activities)	[86]	-	132	Reproducibility: ICC= 0.72 Responsiveness Sensitivity = 0.85%, specificity= 0.84%
WHO-5 well-being index	[87]	-	92	Internal consistency: Cronbach's α = 0.883 Concurrent validity: Spearman correlation with OHIP-49, r = 0.705. Significant association.

References

1. Leeuw Rd. American Academy of Orofacial Pain Guidelines for Assessment, Diagnosis, and Management. 2013.
2. Durham J, Aggarwal V, Davies S, et al. Tempromandibular disorders (TMDs): an update and mangement guidance for primary care from the UK specialist interest group in rorfacial pain and TMDs (USOT). *Royal College of Surgeons England.* 2013.
3. Murphy MK, MacBarb RF, Wong ME, Athanasiou KA. Temporomandibular disorders: a review of etiology, clinical management, and tissue engineering strategies. *Int J Oral Maxillofac Implants.* 2013;28(6):e393-414.
4. Durham J. Oral surgery: part 3. Temporomandibular disorders. *Br Dent J.* 2013;215(7):331-337.
5. Stohler CS. Muscle-related temporomandibular disorders. *J Orofac Pain.* 1999;13(4):273-284.
6. Mills SEE, Nicolson KP, Smith BH. Chronic pain: a review of its epidemiology and associated factors in population-based studies. *Br J Anaesth.* 2019;123(2):e273-e283.
7. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain.* 2005;113(1-2):9-19.
8. 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 Groupdagger. *J Oral Facial Pain Headache.* 2014;28(1):6-27.
9. Coulter A, JENKINSON C, BRUSTER S. The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *Int J Qual Health Care.* 2002;14(5):353-358.
10. Jenkinson C, McGee H. *Health Status Measurement: A Brief but Critical Introduction.* 1998.
11. Griggs CL, Schneider JC, Kazis LE, Ryan CM. Patient-reported Outcome Measures: A Stethoscope for the Patient History. *Ann Surg.* 2017;265(6):1066-1067.
12. Coons SJ, Rao S, Keininger DL, Hays RD. A comparative review of generic quality-of-life instruments. *Pharmacoeconomics.* 2000;17(1):13-35.
13. Devlin NJ, Appleby J, Buxton M, et al. Getting the most out of proms: Putting health outcomes at the heart of NHS decision-making. 2010.
14. Kluzek S, Dean B, Wartolowska KA. Patient-reported outcome measures (PROMs) as proof of treatment efficacy. *BMJ Evidence-Based Medicine.* 2022;27(3):153.
15. Emerson Kavchak AJ, Mischke JJ, Lulofs-MacPherson K, Vendrely AM. The psychometric properties of self-report outcome measures in temporomandibular dysfunction. *Phys Ther Rev.* 2014;19(3):174-185.
16. Aguiar ADS, Nogueira Carrer HC, de Lira MR, Martins Silva GZ, Chaves TC. Patient-Reported Outcome Measurements in Temporomandibular Disorders and Headaches: Summary of Measurement Properties and Applicability. *J Clin Med.* 2021;10(17):3823.
17. Aitken RC. Measurement of feelings using visual analogue scales. *Proc R Soc Med.* 1969;62(10):989-993.
18. Joyce CR, Zutshi DW, Hrubes V, Mason RM. Comparison of fixed interval and visual analogue scales for rating chronic pain. *Eur J Clin Pharmacol.* 1975;8(6):415-420.
19. Huskisson EC. Measurement of pain. *Lancet.* 1974;2(7889):1127-1131.
20. Ni Riordain R, Shirlaw P, Alajbeg I, et al. World Workshop on Oral Medicine VI: Patient-reported outcome measures and oral mucosal disease: current status and future direction. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2015;120(2):152-160.e111.
21. Wiriyakijja P, Fedele S, Porter SR, Mercadante V, Ni Riordain R. Patient-reported outcome measures in oral lichen planus: A comprehensive review of the literature with focus on psychometric properties and interpretability. *J Oral Pathol Med.* 2018;47(3):228-239.

22. Ni Riordain R, McCreary C. Patient-reported outcome measures in burning mouth syndrome - a review of the literature. *Oral Dis.* 2013;19(3):230-235.
23. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). 2011;63(S11):S240-S252.
24. Delgado DA, Lambert BS, Boutris N, et al. Validation of Digital Visual Analog Scale Pain Scoring With a Traditional Paper-based Visual Analog Scale in Adults. *J Am Acad Orthop Surg Glob Res Rev.* 2018;2(3):e088.
25. Slade GD, Spencer AJ. Development and evaluation of the Oral Health Impact Profile. *Community Dent Health.* 1994;11(1):3-11.
26. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs). 2018.
27. Durham J, Steele JG, Wassell RW, et al. Creating a patient-based condition-specific outcome measure for Temporomandibular Disorders (TMDs): Oral Health Impact Profile for TMDs (OHIP-TMDs). *J Oral Rehabil.* 2011;38(12):871-883.
28. Yule PL, Durham J, Playford H, et al. OHIP-TMDs: a patient-reported outcome measure for temporomandibular disorders. *Community Dent Oral Epidemiol.* 2015;43(5):461-470.
29. Dijkers M. Measuring quality of life: methodological issues. *Am J Phys Med Rehabil.* 1999;78(3):286-300.
30. Allen PF, McMillan AS, Walshaw D, Locker D. A comparison of the validity of generic- and disease-specific measures in the assessment of oral health-related quality of life. *Community Dent Oral Epidemiol.* 1999;27(5):344-352.
31. Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders: review, criteria, examinations and specifications, critique. *J Craniomandib Disord.* 1992;6(4):301-355.
32. Ferreira NDR, Marto CMM, de Oliveira AT, Rodrigues MJ, DosSantos MF. Development of core outcome sets for clinical trials in temporomandibular disorders: A study protocol. *PLoS One.* 2022;17(4):e0267722.
33. Loh SY, McLeod RWJ, Elhassan HA. Trismus following different treatment modalities for head and neck cancer: a systematic review of subjective measures. *Eur Arch Otorhinolaryngol.* 2017;274(7):2695-2707.
34. Chiarotto A, Viti C, Sulli A, Cutolo M, Testa M, Piscitelli D. Cross-cultural adaptation and validity of the Italian version of the Central Sensitization Inventory. *Musculoskeletal Science and Practice.* 2018;37:20-28.
35. Wang M, He S, Ji P. Validation of the Centrality of Pain Scale in Chinese-Speaking Patients with Painful Temporomandibular Disorders. *Pain Med.* 2019;20(4):840-845.
36. Barbosa TS, Leme MS, Castelo PM, Gaviao MB. Evaluating oral health-related quality of life measure for children and preadolescents with temporomandibular disorder. *Health & Quality of Life Outcomes.* 2011;9:32.
37. La Touche R, Pardo-Montero J, Gil-Martinez A, et al. Craniofacial Pain and Disability Inventory (CF-PDI): Development and Psychometric Validation of a New Questionnaire. *Pain Physician.* 2014;17(1):95-108.
38. Greghi SM, Dos Santos Aguiar A, Bataglion C, Ferracini GN, La Touche R, Chaves TC. Brazilian Portuguese Version of the Craniofacial Pain and Disability Inventory: Cross-Cultural Reliability, Internal Consistency, and Construct and Structural Validity. *J Oral Facial Pain Headache.* 2018;32(4):389-399.
39. Durham J, Steele JG, Breckons M, Story W, Vale L. DEEP Study: does EQ-5D-5L measure the impacts of persistent oro-facial pain? *J Oral Rehabil.* 2015;42(9):643-650.

40. Rodrigues-Bigaton D, de Castro EM, Pires PF. Factor and Rasch analysis of the Fonseca anamnestic index for the diagnosis of myogenous temporomandibular disorder. *Braz J Phys Ther.* 2017;21(2):120-126.
41. Campos JA, Carrascosa AC, Bonafe FS, Maroco J. Severity of temporomandibular disorders in women: validity and reliability of the Fonseca Anamnestic Index. *Brazilian oral research.* 2014;28:16-21.
42. Balik A, Peker K, Ozdemir-Karatas M. Comparisons of measures that evaluate oral and general health quality of life in patients with temporomandibular disorder and chronic pain. *Cranio.* 2019:1-11.
43. 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(3):219-230.
44. Ohrbach R, Granger C, List T, Dworkin S. Preliminary development and validation of the Jaw Functional Limitation Scale. *Community Dent Oral Epidemiol.* 2008;36(3):228-236.
45. Undt G, Murakami K, Clark GT, et al. Cross-cultural adaptation of the JPF-Questionnaire for German-speaking patients with functional temporomandibular joint disorders. *J Craniomaxillofac Surg.* 2006;34(4):226-233.
46. Aggarwal VR, Lunt M, Zakrzewska JM, Macfarlane GJ, Macfarlane TV. Development and validation of the Manchester orofacial pain disability scale. *Community Dent Oral Epidemiol.* 2005;33(2):141-149.
47. Kallas MS, Crosato EM, Biazevic MG, Mori M, Aggarwal VR. Translation and cross-cultural adaptation of the manchester orofacial pain disability scale. *Journal of oral & maxillofacial research.* 2013;3(4):e3.
48. Stegenga B, de Bont LG, de Leeuw R, Boering G. Assessment of mandibular function impairment associated with temporomandibular joint osteoarthrosis and internal derangement. *J Orofac Pain.* 1993;7(2):183-195.
49. Xu L, Cai B, Fang Z. Translation and validation of a Chinese version of the Mandibular Function Impairment Questionnaire. *J Oral Rehabil.* 2016;43(8):608-614.
50. Campos JA, Carrascosa AC, Maroco J. Validity and reliability of the Portuguese version of Mandibular Function Impairment Questionnaire. *J Oral Rehabil.* 2012;39(5):377-383.
51. Andreu Y, Galdon MJ, Dura E, et al. An examination of the psychometric structure of the Multidimensional Pain Inventory in temporomandibular disorder patients: a confirmatory factor analysis. *Head Face Med.* 2006;2:48.
52. Zucoloto ML, Maroco J, Duarte Bonini Campos JA. Psychometric Properties of the Multidimensional Pain Inventory Applied to Brazilian Patients with Orofacial Pain. *J Oral Facial Pain Headache.* 2015;29(4):363-369.
53. Barbosa C, Manso MC, Reis T, Soares T, Gavinha S, Ohrbach R. Cultural equivalence, reliability and utility of the Portuguese version of the Oral Behaviours Checklist. *J Oral Rehabil.* 2018;45(12):924-931.
54. van der Meulen MJ, Lobbezoo F, Aartman IH, Naeije M. Validity of the Oral Behaviours Checklist: correlations between OBC scores and intensity of facial pain. *J Oral Rehabil.* 2014;41(2):115-121.
55. John MT, Patrick DL, Slade GD. The German version of the Oral Health Impact Profile--translation and psychometric properties. *Eur J Oral Sci.* 2002;110(6):425-433.
56. John MT, Miglioretti DL, LeResche L, Koepsell TD, Hujoel P, Micheelis W. German short forms of the Oral Health Impact Profile. *Community Dent Oral Epidemiol.* 2006;34(4):277-288.
57. Larsson P, List T, Lundstrom I, Marcusson A, Ohrbach R. Reliability and validity of a Swedish version of the Oral Health Impact Profile (OHIP-S). *Acta Odontol Scand.* 2004;62(3):147-152.
58. Segu M, Collesano V, Lobbia S, Rezzani C. Cross-cultural validation of a short form of the Oral Health Impact Profile for temporomandibular disorders. *Community Dent Oral Epidemiol.* 2005;33(2):125-130.

59. Van der Meulen MJ, John MT, Naeije M, Lobbezoo F. Developing abbreviated OHIP versions for use with TMD patients. *J Oral Rehabil.* 2012;39(1):18-27.
60. Murray H, Locker D, Mock D, Tenenbaum HC. Pain and the quality of life in patients referred to a craniofacial pain unit. *J Orofac Pain.* 1996;10(4):316-323.
61. He SL, Wang JH. Validation of the Chinese version of the oral health impact profile for TMDs (OHIP-TMDs-C). *Med Oral Patol Oral Cir Bucal.* 2015;20(2):e161-e166.
62. Bush FM, Harkins SW. Pain-related limitation in activities of daily living in patients with chronic orofacial pain: psychometric properties of a disability index. *J Orofac Pain.* 1995;9(1):57-63.
63. Sugisaki M, Kino K, Yoshida N, Ishikawa T, Amagasa T, Haketa T. Development of a new questionnaire to assess pain-related limitations of daily functions in Japanese patients with temporomandibular disorders. *Community Dent Oral Epidemiol.* 2005;33(5):384-395.
64. He SL, Wang JH, Ji P. Validation of the Pain Resilience Scale in Chinese-speaking patients with temporomandibular disorders pain. *J Oral Rehabil.* 2018;45(3):191-197.
65. Flor H, Behle DJ, Birbaumer N. Assessment of pain-related cognitions in chronic pain patients. *Behav Res Ther.* 1993;31(1):63-73.
66. Rener-Sitar K, John MT, Bandyopadhyay D, Howell MJ, Schiffman EL. Exploration of dimensionality and psychometric properties of the Pittsburgh Sleep Quality Index in cases with temporomandibular disorders. *Health and Quality of Life Outcomes.* 2014;12.
67. Streffer ML, Buchi S, Morgeli H, Galli U, Ettlin D. PRISM (pictorial representation of illness and self measure): a novel visual instrument to assess pain and suffering in orofacial pain patients. *J Orofac Pain.* 2009;23(2):140-146.
68. Lima-Verde AC, Pozza DH, Rodrigues L, Velly AM, Guimaraes AS. Cross-Cultural Adaptation and Validation for Portuguese (Brazilian) of the Pictorial Representation of Illness and Self Measure Instrument in Orofacial Pain Patients. *J Orofac Pain.* 2013;27(3):271-275.
69. de Felicio CM, Melchior Mde O, Da Silva MA. Clinical validity of the protocol for multi-professional centers for the determination of signs and symptoms of temporomandibular disorders. Part II. *Cranio.* 2009;27(1):62-67.
70. Dworkin SF, Sherman J, Mancl L, Ohrbach R, LeResche L, Truelove E. Reliability, validity, and clinical utility of the research diagnostic criteria for Temporomandibular Disorders Axis II Scales: depression, non-specific physical symptoms, and graded chronic pain. *J Orofac Pain.* 2002;16(3):207-220.
71. 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(1):48-62.
72. de Lucena LB, Kosminsky M, da Costa LJ, de Goes PS. Validation of the Portuguese version of the RDC/TMD Axis II questionnaire. *Brazilian oral research.* 2006;20(4):312-317.
73. John MT, Hirsch C, Reiber T, Dworkin S. Translating the research diagnostic criteria for temporomandibular disorders into German: evaluation of content and process. *J Orofac Pain.* 2006;20(1):43-52.
74. Khoo SP, Yap AU, Chan YH, Bulgiba AM. Translating the Research Diagnostic Criteria for Temporomandibular Disorders into Malay: evaluation of content and process. *J Orofac Pain.* 2008;22(2):131-138.
75. Cavalcanti RF, Studart LM, Kosminsky M, de Goes PS. Validation of the multimedia version of the RDC/TMD axis II questionnaire in Portuguese. *Journal of applied oral science : revista FOB.* 2010;18(3):231-236.
76. Shedden Mora MC, Weber D, Neff A, Rief W. Biofeedback-based cognitive-behavioral treatment compared with occlusal splint for temporomandibular disorder: a randomized controlled trial. *Clin J Pain.* 2013;29(12):1057-1065.

77. Dias IM, Bastos RR, Alves RT, Leite ICG. Construction and validation of an questionnaire for evaluating self-medication practised by patients with temporomandibular disorders. *J Oral Rehabil.* 2019;46(5):424-432.
78. Deli R, Macrì L, Mannocci A, La Torre G. Measuring Quality of Life in TMD: use of SF-36. *Italian Journal of Public Health.* 2009;6.
79. He S, Wang J. Validation of the Social support and Pain Questionnaire (SPQ) in patients with painful temporomandibular disorders. *J Headache Pain.* 2017;18(1):57.
80. Funch DP, Marshall JR, Gebhardt GP. Assessment of a short scale to measure social support. *Soc Sci Med.* 1986;23(3):337-344.
81. Nixdorf DR, John MT, Wall MM, Fricton JR, Schiffman EL. Psychometric properties of the modified Symptom Severity Index (SSI). *J Oral Rehabil.* 2010;37(1):11-20.
82. Visscher CM, Ohrbach R, van Wijk AJ, Wilkosz M, Naeije M. The Tampa Scale for Kinesiophobia for Temporomandibular Disorders (TSK-TMD). *Pain.* 2010;150(3):492-500.
83. He SL, Wang JH, Ji P. Validation of the Tampa Scale for Kinesiophobia for Temporomandibular Disorders (TSK-TMD) in patients with painful TMD. *J Headache Pain.* 2016;17.
84. Aguiar AS, Bataglion C, Visscher CM, Bevilaqua Grossi D, Chaves TC. Cross-cultural adaptation, reliability and construct validity of the Tampa scale for kinesiophobia for temporomandibular disorders (TSK/TMD-Br) into Brazilian Portuguese. *J Oral Rehabil.* 2017;44(7):500-510.
85. Gonzalez YM, Schiffman E, Gordon SM, et al. Development of a brief and effective temporomandibular disorder pain screening questionnaire: reliability and validity. *J Am Dent Assoc.* 2011;142(10):1183-1191.
86. Rollman A, Naeije M, Visscher CM. The reproducibility and responsiveness of a patient-specific approach: a new instrument in evaluation of treatment of temporomandibular disorders. *J Orofac Pain.* 2010;24(1):101-105.
87. Ismail F, Lange K, Gillig M, et al. WHO-5 well-being index as screening instrument for psychological comorbidity in patients with temporomandibular disorder. *Cranio-the Journal of Craniomandibular Practice.* 2018;36(3):189-194.