Patient-Reported Outcome Measures Used in Temporomandibular Disorders: A Review of the Literature

Dina Taimeh, DDS, MSc, PhD
Department of Maxillofacial Medicine and Surgery
UCL Eastman Dental Institute
University College London, London, United Kingdom;
Department of Oral and Maxillofacial Surgery, Oral Medicine and Periodontology
Faculty of Dentistry
University of Jordan, Amman, Jordan

Richeal Ni Riordain, MBBS, BDS, MA (Higher Ed), PhD
Department of Maxillofacial Medicine and Surgery
UCL Eastman Dental Institute
University College London;
Oral Medicine Unit
Cork University Dental School and Hospital, Cork, Ireland

Stefano Fedele, DDS, PhD
Department of Maxillofacial Medicine and Surgery
UCL Eastman Dental Institute
University College London;
Oral Medicine Unit
Eastman Dental Hospital
University College London Hospitals Trust;
NIHR University College London Hospitals Biomedical Research Centre, London, United Kingdom

Rachel Leeson, BDS, MSc
Department of Maxillofacial Medicine and Surgery
UCL Eastman Dental Institute
University College London, London, United Kingdom

Correspondence to:
Dr Dina Taimeh
UCL Eastman Dental Institute, University College London, Rockefeller Building
20 University Street, London WC1E 6DE, UK.
Email: dina.taimeh.17@ucl.ac.uk

Submitted June 1, 2022; accepted February 20, 2023.

Aims: To identify the range of patient-reported outcome measures (PROMs) used in TMD studies, summarize the available evidence for their psychometric properties, and provide guidance for the selection of such measures. **Methods:** A comprehensive search was conducted to retrieve articles published between 2009 and 2018 containing a patient-reported measure of the effects of TMDs. Three databases were searched: MEDLINE, Embase, and Web of Science. **Results:** A total of 517 articles containing at least one PROM were included in the review, and 57 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; PROMs describing psychologic status; and PROMs describing quality of life and general health. The most commonly used PROM was the visual analog scale. However, a wide range of verbal descriptors was employed. The Oral Health Impact Profile-14 and Beck Depression Inventory were the most commonly used PROMs describing the effect of TMDs on quality of life and psychologic status, respectively. Additionally, the Oral Health Impact Profile (various versions) and the Research Diagnostic Criteria Axis II questionnaires were the instruments most repeatedly tested in a TMD population, and these instruments have undergone cross-cultural validation in several languages. **Conclusion:** A wide range of PROMs have been used to describe the impact of TMDs on patients. Such variability may limit the ability of researchers and clinicians to evaluate the efficacy of different treatments and make meaningful comparisons. *J Oral Facial Pain Headache 2023;37:113–129. doi: 10.11607/ofph.3264*

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

Temporomandibular disorders (TMDs) is a collective term embracing a number of clinical conditions that involve the masticatory musculature, the temporomandibular joint (TMJ), and associated structures.[1] This condition is the most common cause of chronic pain in the facial region.[2] Associated symptoms include pain, restricted mouth opening, deviation in mandibular movements, clicking noises of the joint, headache in the temporal region, and psychologic effects.[3,4] Among the various categories grouped under this umbrella term, muscle problems are the largest.[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.[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 psychologic and emotional functioning associated with chronic pain.

Traditionally, health care has been assessed in terms of the technical and physiologic outcomes of treatment.[9] In more recent times, however, health care organizations are striving to achieve services that are not only clinically effective and evidence based but also beneficial to and effective for patients as judged from their own perspectives.[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 clinical decision-making, and enhance communication between patients and physicians.[11] Many instruments exist that measure the intensity of pain, quality of life, psychologic distress, and disability. Some are generic for use in a wide range of conditions and settings, and some are condition specific, 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 to provide guidance for the selection of such measures.

Materials and Methods

Search Strategy

A comprehensive search was conducted in January 2019 to retrieve published articles concerned with patient-reported assessment of the effects of TMDs. 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 to 2018. Articles assessing the psychometric properties of PROMs were located using the same search strategy; however, no time restrictions were applied (ie, all the articles yielded from the search strategy were screened up to January 2019).

The included studies were clinical trials and observational studies of TMDs (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 radiologic outcomes only, studies containing PROMs that reported on the side effects after a specific intervention (eg, complications of surgery), systematic and literature reviews, case reports, book chapters, conference proceedings, commentary or author opinions, animal studies, and studies with an unavailable full text.

Data Extraction

A study-specific Microsoft 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 TMDs, classification system used, the PROM used, and the follow-up time point (if any). Additional data were also extracted from studies assessing 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 databases yielded 3,452 articles in total. After applying the exclusion criteria, 517 articles containing at least one PROM remained. An additional 57 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).

PROMs Commonly Used in TMDs

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 psychologic status and satisfaction; and PROMs describing quality of life and general health. Table 1 shows the identified PROMs and their frequency of use.

The most commonly used PROM was the visual analog scale (VAS), with 59.5% of the trials using this instrument. Various verbal descriptors were employed, including 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 100-mm scale (or a 10-cm scale). In a few studies, however, the VAS was associated with scales ranging from 0 to 3, 0 to 4, 0 to 5, 0 to 6, and −5 to 5 (highlighting possible misuse of the VAS).

Likert point and numeric rating scales (NRS) 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 scales 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 to 10, 0 to 3, 0 to 5, 0 to 6, and 1 to 4. Among the other

Table 1 PROMS Identified and Their Frequency of Use in the 517 Studies Examined

Name of PROM	Frequency of use (no. of studies)
Severity of symptoms and improvement	
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 (NDI)	6
Characteristic Pain Intensity (CPI)	4
3Q/TMD	3
Brief Symptom Inventory (BSI)	3
The Pain-Related Self-Statements Scale	3
Chronic pain grade	2
Color analog 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
Oral Behavior Checklist (OBC)	2
Tinnitus Handicap Inventory (THI)	2
Widespread Pain Index (WPI)	2
Craniofacial Pain and Disability Inventory (CF-PDI)	1
Food Intake Ability (FIA) Index	1
Mann Assessment of Swallowing Ability (MASA) score	1
Pictorial Representation of Illness and Self-Measure (PRISM)	1
Screening for Somatoform Symptoms (SOMS-7)	1
Symptom Interference Questionnaire—Revised (SIQR)	1
The Battery for Health Improvement	1
Gracely Pain Scale	1
Jaw Pain and Function (JPF)-Questionnaire	1
Pain Behavior Questionnaire	1
Patient-Specific Functional Scale (PSFS)	1
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 (PSQI)	17
Short Form-36 Health Survey (SF-36)	14
Oral Health Impact Profile (OHIP)	9
General Health Questionnaire-7 (GHQ-7)	7
Short-Form 12 Health Survey (SF-12)	5
Epworth Sleepiness Scale (ESS)	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

Table 1 PROMS Identified and Their Frequency of Use in the 517 Studies Examined *(continued)*

Name of PROM	Frequency of use (no. of studies)
Child Perceptions 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
Flanagan Quality of Life Scale	1
Life Experiences Survey (LES)	1
Multidimensional Fatigue Inventory—Short Form	1
Sleep Assessment Questionnaire (SAQ)	1
Psychologic 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
Generalized 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
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

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 it incurred. A few PROMs described other symptoms associated with TMDs, such as the Neck Disability Index (NDI; n = 6), Tinnitus Handicap Inventory (THI; n = 2), Headache Impact Test-6 (HIT-6; n = 2), and Food Intake Ability (FIA) index (n = 1).

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

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

Psychometric Properties of PROMs Used in TMDs

Several PROMs identified in this 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 OHIP (several versions) were the instruments most repeatedly tested in a TMD population, and these instruments have undergone cross-cultural validation into several languages. The search also identified a TMD-specific variant of OHIP, the OHIP-TMD. The reported psychometric properties were internal consistency (Cronbach $\alpha = 0.94$), 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 2 items relevant to facial pain patients (Cronbach $\alpha = 0.97$).

Discussion

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

Emerson et al provided an assessment of the psychometric properties of some tools in a TMD population in 2014.[15] This 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 with complete psychometric properties established. Aguiar et al also examined the psychometric properties of 10 common condition-specific PROMs and had similar conclusions, noting the need for further studies on psychometric properties.[16]

In the present review of 517 studies, 106 PROMs used to assess the effects of TMD on patients were identified, and an additional 57 papers that tested the psychometric properties of some tools in a TMD population, including cross-cultural adaptation, were identified. The most used PROM was the VAS. The pain VAS mimics the continuous VAS 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 the 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 rationalized 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 pooling data. The VAS may also have been misnamed in a small number of studies where different increments were utilized (eg, 0 to 3). The VAS is usually displayed as a 10-cm line that represents a continuum between the two ends of the scale.[24] Hence, the scores on such scales may be better labeled as Likert point scales.

The most frequently used oral health quality of life PROM in this review was the OHIP-14. The items for the OHIP were generated following interviews with patients from private dental practices, primary care clinics, and prosthetic clinics in a dental hospital[25]; therefore, it may not be specific enough for patients with TMDs 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 the 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 psychologic status PROMs. Several have been created to describe the symptoms of TMDs, such as the 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 studies have used generic PROMs to describe the quality of life and psychologic status of the patients. Condition-specific PROMs are more sensitive and have greater discriminatory ability to detect small changes over time.[29,30]

Numerous PROMs are used to describe the impact of TMDs on patients. Such variability may limit the ability of researchers and clinicians to evaluate the efficacy of different treatments, pool data, and make 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 RDC/TMD has also been proposed to provide a comprehensive diagnostic and classification system for the subtypes of TMDs. These criteria, first proposed in 1992,[31] were 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 behavior, psychologic status, and psychosocial functioning.[8] The present 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 radiologic outcomes; therefore, a comprehensive evaluation of psychosocial functioning may not be crucial to the researchers. It is, however, also important to mention that a set of core outcomes for clinical trials in TMDs is currently under development.[32]

It was also noted that while some studies employed subjective measures to assess function, other studies—and sometimes 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 limitations in 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 patient-reported measures are essential in the assessment of TMDs. Hence, TMD classification criteria, for example the DC/TMD, utilize the outcome of the clinical exam in addition to a patient-completed symptom questionnaire to establish a diagnosis. Additionally, Loh et al reported in their systematic review on trismus instruments that the correlation between subjective and objective measures was strong overall, and the findings of some studies that 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 databases. Indeed, the results of the search might be different if studies in other languages were included or the search expanded to other databases with no time restrictions.

A comprehensive search was carried out to locate papers testing the psychometric properties of the different tools in a TMD population. However, the present paper 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 is presented in Table 2[34–87] to enable readers to judge the suitability of each PROM for their own setting.

Conclusions

Condition-specific PROMs to assess the psychologic 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 for describing pain intensity, oral health–related quality of life, and psychologic status of the participants, respectively, in TMD studies.
- 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.

Acknowledgments

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 D.T. The authors declare no conflicts of interest.

References

1. de Leeuw Rd, Klasser GD, eds. American Academy of Orofacial Pain Guidelines for Assessment, Diagnosis, and Management, ed 5. Quintessence, 2013.
2. Durham J, Aggarwal V, Davies S, Harrison SD. Temporomandibular disorders (TMDs): An update and management guidance for primary care from the UK Specialist Interest Group in Orofacial Pain and TMDs (USOT). Royal College of Surgeons Faculty of Dental Surgery, 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:e393–e414.
4. Durham J. Oral surgery: Part 3. Temporomandibular disorders. Br Dent J 2013;215:331–337.
5. Stohler CS. Muscle-related temporomandibular disorders. J Orofac Pain 1999;13:273–284.
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:e273–e283.
7. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113:9–19.
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 group. J Oral Facial Pain Headache 2014;28:6–27.
9. Jenkinson C, Coulter A, 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;14:353–358.
10. Jenkinson C, McGee HM. Health Status Measurement: A Brief But Critical Introduction. Radcliffe, 1998.
11. Griggs CL, Schneider JC, Kazis LE, Ryan CM. Patient-reported outcome measures: A stethoscope for the patient history. Ann Surg 2017;265:1066–1067.
12. Coons SJ, Rao S, Keininger DL, Hays RD. A comparative review of generic quality-of-life instruments. Pharmacoeconomics 2000;17:13–35.
13. Devlin NJ, Appleby J. Getting the most out of PROMs: Putting health outcomes at the heart of NHS decision-making. The King's Fund, 2010.
14. Kluzek S, Dean B, Wartolowska KA. Patient-reported outcome measures (PROMs) as proof of treatment efficacy. BMJ Evid Based Med 2022;27:153–155.
15. Emerson AJ, MacPherson KL, Mischke JJ, Vendrely A. The psychometric properties of self-report outcome measures in temporomandibular dysfunction. Phys Ther Rev 2014;19:174–185.
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;10:3823.
17. Aitken RC. Measurement of feelings using visual analogue scales. Proc R Soc Med 1969;62:989–993.
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;8:415–420.
19. Huskisson EC. Measurement of pain. Lancet. 1974;2(7889): 1127–1131.
20. Ní Ríordáin 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;120:152–160.e11.
21. Wiriyakijja P, Fedele S, Porter SR, Mercadante V, Ní Ríordáin 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;47:228–239.
22. Ní Ríordáin R, McCreary C. Patient-reported outcome measures in burning mouth syndrome—A review of the literature. Oral Dis 2013;19:230–235.
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). Arthritis Care Res (Hoboken) 2011;63(suppl 11):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:e088.
25. Slade GD, Spencer AJ. Development and evaluation of the oral health impact profile. Community Dent Health 1994;11:3–11.
26. Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). COSMIN Manual for Systematic Reviews of PROMs. February 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;38:871–883.
28. Yule PL, Durham J, Playford H, et al. OHIP-TMDs: A patient-reported outcome measure for temporomandibular disorders. Community Dent Oral Epidemiol 2015;43:461–470.
29. Dijkers M. Measuring quality of life: Methodological issues. Am J Phys Med Rehabil 1999;78:286–300.
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;27:344–352.
31. Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders: Review, criteria, examinations and specifications, critique. J Craniomandib Disord 1992;6:301–355.
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: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;274:2695–2707.
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. Musculoskelet Sci Pract 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;20:840–845.
36. Barbosa TS, Leme MS, Castelo PM, Gavião MB. Evaluating oral health-related quality of life measure for children and preadolescents with temporomandibular disorder. Health Qual Life Outcomes 2011;9:32.
37. La Touche R, Pardo-Montero J, Gil-Martínez A, et al. Craniofacial pain and disability inventory (CF-PDI): Development and psychometric validation of a new questionnaire. Pain Physician 2014;17:95–108.
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;32:389–399.
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;42:643–650.

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;21:120–126.
41. Campos JADB, Carrascosa AC, Bonafé FSS, Maroco J. Severity of temporomandibular disorders in women: Validity and reliability of the Fonseca anamnestic index. Braz Oral Res 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;39:310–320.
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;22:219–230.
44. Ohrbach R, Granger C, List T, Dworkin S. Preliminary development and validation of the jaw functional limitation scale. Community Dent Oral Epidemiol 2008;36:228–236.
45. Undt G, Murakami K-I, Clark GT, et al. Cross-cultural adaptation of the JPF-Questionnaire for German-speaking patients with functional temporomandibular joint disorders. J Craniomaxillofac Surg 2006;34:226–233.
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;33:141–149.
47. Kallás MS, Crosato EM, Biazevic MGH, Mori M, Aggarwal VR. Translation and cross-cultural adaptation of the manchester orofacial pain disability scale. J Oral Maxillofac Res 2013;3: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;7:183–195.
49. Xu L, Cai B, Fang Z. Translation and validation of a Chinese version of the mandibular function impairment questionnaire. J Oral Rehabil 2016;43:608–614.
50. Campos JADB, Carrascosa AC, Maroco J. Validity and reliability of the Portuguese version of mandibular function impairment questionnaire. J Oral Rehabil 2012;39:377–383.
51. Andreu Y, Galdon MJ, Durá 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;29:363–369.
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;45:924–931.
54. van der Meulen MJ, Lobbezoo F, Aartman IHA, Naeije M. Validity of the oral behaviours checklist: Correlations between OBC scores and intensity of facial pain. J Oral Rehabil 2014;41:115–121.
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;110:425–433.
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;34:277–288.
57. Larsson P, List T, Lundström I, Marcusson A, Ohrbach R. Reliability and validity of a Swedish version of the oral health impact profile (OHIP-S). Acta Odontol Scand 2004;62:147–152.
58. Segù 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;33:125–130.
59. van der Meulen MJ, John MT, Naeije M, Lobbezoo F. Developing abbreviated OHIP versions for use with TMD patients. J Oral Rehabil 2012;39:18–27.
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;10:316–323.
61. He S-L, Wang J-H. Validation of the Chinese version of the oral health impact profile for TMDs (OHIP-TMDs-C). Med Oral Patol Oral Cir Bucal 2015;20: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;9:57–63.
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;33:384–395.
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;45:191–197.
65. Flor H, Behle DJ, Birbaumer N. Assessment of pain-related cognitions in chronic pain patients. Behav Res Ther 1993;31:63–73.
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 Qual Life Outcomes 2014;12:10.
67. Streffer M-L, Büchi S, Mörgeli 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;23:140–146.
68. Lima-Verde AC, Pozza DH, Rodrigues LLFR, Velly AM, Guimarães 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;27:271–275.
69. de Felicio CM, Melchior MdO, Da Silva MAM. Clinical validity of the protocol for multi-professional centers for the determination of signs and symptoms of temporomandibular disorders. Part II. Cranio 2009;27:62–67.
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;16:207–220.
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;24:48–62.
72. de Lucena LBS, Kosminsky M, da Costa LJ, de Góes PS. Validation of the Portuguese version of the RDC/TMD axis II questionnaire. Braz Oral Res 2006;20:312–317.
73. John MT, Hirsch C, Reiber T, Dworkin SF. Translating the research diagnostic criteria for temporomandibular disorders into German: Evaluation of content and process. J Orofac Pain 2006;20:43–52.
74. Khoo S-P, Yap AUJ, Chan YH, Bulgiba AM. Translating the research diagnostic criteria for temporomandibular disorders into Malay: Evaluation of content and process. J Orofac Pain 2008;22:131–138.
75. Cavalcanti RF, Studart LM, Kosminsky M, de Goes PSA. Validation of the multimedia version of the RDC/TMD axis II questionnaire in Portuguese. J Appl Oral 2010;18:231–236.
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;29:1057–1065.

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;46:424–432.
78. Deli R, Macrì LA, Mannocci A, La Torre G. Measuring quality of life in TMD: Use of SF-36. Italian J Public Health 2009;6. doi: https://doi.org/10.2427/5790
79. He S, Wang J. Validation of the social support and pain questionnaire (SPQ) in patients with painful temporomandibular disorders. J Headache Pain 2017;18:57.
80. Funch DP, Marshall JR, Gebhardt GP. Assessment of a short scale to measure social support. Soc Sci Med 1986;23:337–344.
81. Nixdorf DR, John MT, Wall MM, Fricton JR, Schiffman EL. Psychometric properties of the modified symptom severity index (SSI). J Oral Rehabil 2010;37:11–20.
82. Visscher CM, Ohrbach R, van Wijk AJ, Wilkosz M, Naeije M. The Tampa scale for kinesiophobia for temporomandibular disorders (TSK-TMD). Pain 2010;150:492–500.
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;17:109.
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;44:500–510.
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;142:1183–1191.
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;24:101–105.
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 2018;36:189–194.

Table 2 Details of the Psychometric Properties of some PROMs in a TMD Population

PROM	Study	Domains/factors	No. of TMD patients	Psychometric testing
Central Sensitization Inventory (CSI)-Italian	Chiarotto et al, 2018[34]	–	37	Cross-cultural adaptation Structural validity: EFA Construct validity: Pearson correlation with 11-point NRS for pain intensity = 0.427, SF-36 = −0.479, HADS = 0.706, PSEQ = −0.618. All have significant correlations. Internal consistency: Cronbach α = 0.87
Centrality of Pain Scale-Chinese	Wang et al, 2019[35]	–	166	Cross-cultural adaptation Internal consistency: Cronbach α = 0.942 Test-retest (30 patients, 1 wk): ICC = 0.815–0.929 Construct validity: EFA, 1 factor Convergent validity: Pearson correlation with: Catastrophizing Scale (r = 0.57) and PSEQ (r = −0.42). Both have significant correlation.
Child Perception Questionnaire (CPQ) 8–10 y CPQ 11–14 y	Barbosa et al, 2011[36]	-Oral symptoms -Functional limitations -Emotional well-being -Social well-being	547	Criterion validity: Spearman correlation with pain scores (obtained from question 3 of the RDC/TMD Axis II). CPQ 8–10: r = 0.18/nonsignificant, CPQ 11–14: r = 0.32/significant Discriminant construct validity Correlational construct validity Internal reliability (internal consistency) CPQ 8–10: Cronbach α = 0.93, CPQ 11–14: Cronbach α = 0.94
Craniofacial Pain and Disability Inventory (CFPDI)-Spanish	La Touche et al, 2014[37]	-Pain and disability -Jaw functional status	192	Test-retest reliability (106 patients, 12 d): ICC = 0.90 Internal consistency: Cronbach α = 0.88 Construct validity: EFA, 2 factors Responsiveness: SEM = 2.4 Convergent validity: Pearson 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	Greghi et al, 2018[38]	–	100	Cross-cultural adaptation Internal consistency: Cronbach α = 0.77–0.86 Construct validity: Pearson 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: CFA, 3 factors Test-retest (60 patients, 1 wk): ICC = 0.97
EQ-5D-5L	Durham et al, 2015[39]	-Mobility -Self-care -Usual activities -Pain/discomfort -Anxiety/ depression	66	Convergent validity Spearman rho with MPI for each subscale.
Fonseca Anamnestic Index (FAI)	Rodrigues-Bigaton et al, 2017[40]	–	94	Structural validity: EFA Overall correlation between items: Spearman correlation: Some items showed good correlation, but not all items were correlated, suggesting more than one dimension in the FAI. Internal consistency: Cronbach α = 0.7. Rasch analysis
	Campos et al, 2014[41]	–	700 (normal population, assuming 40% are TMD)	CFA Convergent validity: AVE = 0.513, CR = 0.878 Internal consistency: Cronbach α = 0.745. Reproducibility (62 patients, 1 wk): kappa = 0.89. Concurrent validity: Correlation analysis with MFIQ: r = 0.66 (Questions 8 and 10 were below the adequate values and were thus excluded from the original model.)

Table 2 Details of the Psychometric Properties of some PROMs in a TMD Population *(continued)*

PROM	Study	Domains/factors	No. of TMD patients	Psychometric testing
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	Balik et al, 2019[42]	SCL-90-R: -Somatization -Obsessive-compulsive interpersonal sensibility -Depression -Anxiety -Anger-hostility -Phobic-anxiety -Paranoid ideation -Psychoticism -Other items	104	Internal consistency: Cronbach α JDC = 0.76 CPI = 0.79 SCL-90-R somatization = 0.87 SCL-90-R depression = 0.93 OHIP-14 = 0.86 SF-36 physical health = 0.83 SF-36 mental health = 0.82
JFLS-20, 8	Ohrbach et al, 2008[43]	-Mastication -Vertical jaw mobility -Emotional and verbal expression	31	Fitness of model/item reliability Rasch methodology Temporal stability (1–2 wk): CCC JFLS-20 = 0.87, JFLS-8 = 0.81 Internal consistency: Cronbach α JFLS-20 = 0.95, JFLS-8 = 0.87 Correlation of subscales: JFLS-20 = 0.9422
	Ohrbach et al, 2008[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 CPI = 0.49, and STAI = 0.17
Jaw Pain and Function (JPF)-German	Undt et al, 2006[45]	-Jaw pain -Jaw function	137	Cross-cultural adaptation Concurrent construct validity (97 patients) Pearson correlation with maximum interincisal distance, $r = -0.213$. Significant correlation. Test-retest reliability (40 patients, 1 d and 1 wk) Pearson correlation, 1 d: $r = 0.91$, 1 wk: $r = 0.93$. Both are significant. Internal consistency of verbal subscales Cronbach α: Pain score = 0.85, ADL score = 0.94, function score = 0.68.
Manchester Orofacial Pain Disability Scale	Aggarwal et al, 2005[46]	-Physical -Psychologic	171	Internal consistency Cronbach 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
	Kallás et al, 2013[47]	–	50	Cross-cultural validation Internal consistency: Cronbach α = 0.9 Test-retest reproducibility 15–20 d: ICC = 0.924 Criterion validity: Correlations with OHIP-14 ($r = 0.857$) and VAS for pain intensity ($r = 0.758$). Both are significant. Interitem correlation Factor analysis: 2 factors
Mandibular Function Impairment Questionnaire (MFIQ)	Stegenga et al, 1993[48]	-Functional capacity -Feeding	95	Convergent validity Internal consistency: Cronbach α = 0.63 to 0.95. Factor analysis: 3 factors.
MFIQ-Chinese	Xu et al, 2016[49]	–	352	Cross-cultural adaptation Internal consistency: Cronbach α for factor 1: 0.925, for factor 2 = 0.72. Test-retest reliability (78 patients–7 d): ICC for factor 1 = 0.895, for factor 2 = 0.720. Content validity: evaluated by 20 dentists and 5 physical therapists. Construct validity: EFA and CFA, 2 factors. Face validity: consensus among 8 specialists.

Table 2 Details of the Psychometric Properties of some PROMs in a TMD Population *(continued)*

PROM	Study	Domains/factors	No. of TMD patients	Psychometric testing
MFIQ-Portuguese	Campos et al, 2012[50]	–	249	Factorial validity: CFA, 2 factors Internal consistency: Cronbach α for functional capacity dimension = 0.874, for feeding dimension = 0.918. Intrarater reproducibility (62 patients, 1 wk): ICC for functional capacity dimension 1 = 0.895, for feeding dimension = 0.825. Temporal stability (test-retest reliability): Pearson correlation for dimension 1 *r* = 896, for dimension 2 *r* = 0.826. Face validity: evaluated by 6 dentistry professionals (specialists on TMDs) and 3 experts of the English language. Content validity: assessed by 21 dentists with expertise in TMDs. Convergent and discriminant validity were assessed, respectively, by AVE, CR, and bivariate correlations between factors.
Multidimensional Pain Inventory (MPI)-Spanish	Andreu et al, 2006[51]	-Pain impact -Responses by significant others -Activities	114	Cross-cultural adaptation Internal consistency: Cronbach α > 0.7 for all items. CFA.
MPI-Brazilian	Zucoloto et al, 2015[52]	–	31	Convergent validity: AVE and CR Internal consistency: Cronbach α = 0.80–0.94 CVR: 15 experts in the field of dentistry. Construct validity: CFA
OBC-Portuguese	Barbosa et al, 2018[53]	-Activities during sleep -Activities during waking hours	120	Cultural adaptation Test-retest reliability (120 patients, 2 wk) ICC = 0.998. Temporal stability: weighted kappa > 0.946 Item agreement between English and Portuguese OBC: weighted kappa > 0.934 Internal consistency: Cronbach α = 0.64 Convergent and discriminant validity
OBC-Dutch	van der Meulen et al, 2014[54]	–	155	Cross-cultural validity Test-retest reliability (35 patients, 2 wk): ICC = 0.86 Concurrent validity: Spearman correlation with Dutch Oral Parafunctions Questionnaire *r* = 0.757, RDC-CPI *r* = 0.069, Dutch SCL-90 depression *r* = 0.485, somatization *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: 0.389–0.892
OHIP-49–German	John et al, 2002[55]	-Functional limitation -Physical pain -Psychologic discomfort -Physical disability -Psychologic disability -Social disability -Handicap	67	Cross-cultural validation Groups validity: Point-biserial correlations Responsiveness (1 mo): Effect size calculation with paired *t* test.
OHIP-5, 14, 21–German	John et al, 2006[56]	–	175	Validity and internal consistency: Cronbach α = 0.65–0.92 Responsiveness: standardized effect size = 0.55–0.95 Construct validity: Point-biserial correlations
OHIP-49–Swedish	Larsson et al, 2004[57]	–	30	Test-retest reliability: ICC = 0.87–0.98. Construct validity: Spearman correlation with JFLS (*r* = 0.76), SCL-90 (*r* = 0.65), self-reported health (*r* = 0.61). Internal reliability: Cronbach α = 0.83–0.91

Table 2 Details of the Psychometric Properties of some PROMs in a TMD Population *(continued)*

PROM	Study	Domains/factors	No. of TMD patients	Psychometric testing
OHIP-Italian	Segù et al, 2005[58]	–	124	Cross-cultural validation Content validation: group of experts Internal consistency: Cronbach α = 0.71–0.86 Construct validation: known-groups analysis. Criterion-related validation EFA, 7 factors
OHIP-5, 14, 48–Dutch	van der Meulen et al, 2012[59]	–	245	Internal consistency: Cronbach α OHIP-48 = 0.96, OHIP-14 = 0.9, OHIP-5 = 0.67 Test-retest reliability (64 patients, 2 wk): ICC OHIP-48 = 0.82, OHIP-14 = 0.8, OHIP-5 = 0.69 Construct validity Convergent validity: Spearman rho with pain-related disability score: OHIP-48 = 0.46, OHIP-14 = 0.46, OHIP-5 = 0.39, self-reported oral health status OHIP-48 = 0.28, OHIP-14 = 0.19, OHIP-5 = 0.21. Group validity: *t* test between patients with and without complaints and Spearman rho (with CPI and biting activities)
OHIP-30–OFP	Murray et al, 1996[60]	–	121	Internal consistency: Cronbach α = 0.97.
OHIP-TMD	Durham et al, 2011[27]	–	110	Convergent validity: Spearman rho correlation with MPI = 0.751, VAS = 0.576, without the 2 new items. Both are significant. Internal consistency: Cronbach α = 0.942 (without the 2 new items).
OHIP-TMD	Yule et al, 2015[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-TMD vs OHIP-49) Test-retest reliability: ICC = 0.805 Internal consistency: Cronbach α = 0.95 at baseline, 0.96 at follow-up.
OHIP-TMD–Chinese	He and Wang, 2015[61]	–	156	Cross-cultural validation Internal consistency Cronbach α = 0.917. Test-retest reliability (30 patients, 2 wk): ICC = 0.899. Structural validity: Factor analysis, 5 factors. Convergent validity: Global rating of oral health question = 0.548. Significant correlation.
Pain Disability Index	Bush and Harkins, 1995[62]	-Family/home responsibilities -Recreation -Social activity -Occupation -Sexual behavior -Self-care -Life support activities	197	Factor structure.
Pain-related Limitations of Daily Functions (LDF-TMD-Q)-Japanese	Sugisaki et al, 2005[63]	-Limitations in executing a certain task -Limitations in mouth opening -Limitations in sleeping	456	Factor validity EFA, 3 factors CFA 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 α = 0.81, and split-half estimation (Guttmann method) $r = 0.76$ ($P < .05$)

Table 2 Details of the Psychometric Properties of some PROMs in a TMD Population *(continued)*

PROM	Study	Domains/factors	No. of TMD patients	Psychometric testing
Pain Resilience Scale-Chinese	He et al, 2018[64]	-Cognitive/affective positivity -Behavioral perseverance	152	Cross-cultural validity CFA, 2 factors Internal consistency: Cronbach α = 0.92 Test retest reliability (30 patients, 2 wk): ICC = 0.92 Convergent validity: Spearman correlation with Connor-Davidson Resilience Scale = 0.61 to 0.65 and TSK-TMD = −0.46 to −0.41
Pain-Related Control Scale (PRCS)	Flor et al, 1993[65]	-Helplessness -Resourcefulness	44	Internal consistency: Cronbach 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)	Flor et al, 1993[65]	-Catastrophizing -Coping	44	Internal consistency: Cronbach α = 0.92, 0.88. Convergent Discriminant validity Factor analysis Stability (test-retest reliability): PRSS Catastrophizing = 0.87, PRSS Coping = 0.77
Pittsburgh Sleep Quality Index (PSQI)	Rener-Sitar et al, 2014[66]	-Subjective sleep quality -Sleep latency -Sleep duration -Habitual sleep efficiency - Sleep disturbances -Use of sleeping medication -Daytime dysfunction	609	EFA, 1 factor Model fit: CFA Internal consistency: Cronbach α = 0.75 Interitem correlation: Pearson correlation coefficient = 0.3 Test-retest reliability: ICC = 0.86 Convergent validity: Spearman rho coefficient with questions from the GHQ, Q1 = 0.43, Q2 = 0.48
PRISM (pictorial representation of illness and self-measure)	Streffer et al, 2009[67]	–	70	Construct validity: Pearson 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)	Lima-Verde et al, 2013[68]	–	42	Cross-cultural translation Content validity: Pearson correlations with 0–10 NPS (moderate −0.42), Insomnia Severity Index (wk −0.24), HADS-A (wk −0.25), HADS-D (wk −0.22) Temporal stability (30 patients, 3 d): ICC = 0.991
ProTMDMulti	de Felicio et al, 2009[69]	–	30	Criterion validity: Spearman R with Helkimo index = 0.65. Significant correlation. Construct validity: Comparison results between pre- and posttreatment and between TMD group and control group
RDC/TMD Axis II	Dworkin et al, 2002[70]	Graded chronic pain, depression, somatization with and without pain, JDC	362	Concurrent validity of SCL-90 depression: Pearson correlations with BDI = 0.69, Centre for Epidemiologic Studies for Depression = 0.78 Internal consistency reliability: Cronbach α SCL-90 = 0.91, nonspecific physical symptoms = 0.82, GCPS = 0.71 Construct validity of the nonspecific physical symptoms scale: EFA, 2 factors Clinical utility: Sensitivity = 0.91, specificity = 0.41.

Table 2 Details of the Psychometric Properties of some PROMs in a TMD Population *(continued)*

PROM	Study	Domains/factors	No. of TMD patients	Psychometric testing
RDC/TMD Axis II	Ohrbach et al, 2010[71]	–	626	Internal consistency: Cronbach 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 CCC with CESD = 0.85 and with SF-12 = −0.70 SCL-90 nonspecific 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 wk): SCL-90 depression CCC = 0.63–0.78. SCL-90 nonspecific physical symptoms CCC = 0.63–0.78 GCPS-CPI (3 d): CCC = 0.91 GCPS activity interference: CCC = 0.89 GCPS chronic pain grade: weighted kappa = 0.87 Discriminant validity: Lin CCC with MPI. Criterion validity Clinical utility of the depression instrument by calculating PPV, NPV
RDC/TMD Axis II-Portuguese	de Lucena et al, 2006[72]	–	55	Internal consistency: Cronbach α = 0.72 Reliability: kappa = 0.73 to 0.9 Test retest reliability (45 patients, 2 wk): Cohen kappa scale/Axis 1. Spearman rank correlation = 0.727–0.821 Concurring validation: Spearman correlation with Oral Impacts on Daily Performances = 0.306–0.602, OHIP-14 = 0.336–0.598.
RDC/TMD Axis II-German	John et al, 2006[73]	–	378	Cross-cultural adaptation Test-retest reliability (27 patients, 1 to 2 wk): ICC jaw disability list = 0.76, GCPS = 0.92. Internal consistency: Cronbach α jaw disability list = 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	Khoo et al, 2008[74]	–	40	Cross-cultural validity Internal consistency: Cronbach α GCPS = 0.77, nonspecific physical symptoms = 0.71, depression = 0.88. Test-retest reliability (40 patients, 1 wk): ICC GCPS = 0.97, nonspecific physical symptoms = 0.94, depression = 0.95 Discriminant validity: *t* test of means between patients with pain symptoms and without symptoms SEM
Multimedia Version of the RDC/TMD Axis II-Portuguese	Cavalcanti et al, 2010[75]	–	30	Internal consistency: Cronbach α = 0.94 Convergent validity: Spearman rank correlation Reproducibility (1 d): Spearman rank correlation test = 0.670–0.913
Screening for Somatoform Symptoms (SOMS-7)	Shedden Mora et al, 2013[76]	-Somatization symptom count -Somatization severity index	58	Internal consistency: Cronbach α = 0.88
Self-medication questionnaire	Dias et al, 2019[77]	-Symptoms -Opinion about self-medication	110	Face validity (content validity): interviews with 10 patients and expert opinion. Internal reliability: Cronbach α = 0.844. EFA, 2 factors. Reproducibility (11 patients, 15 d): weighted kappa coefficient = 0.81.

Table 2 Details of the Psychometric Properties of some PROMs in a TMD Population *(continued)*

PROM	Study	Domains/factors	No. of TMD patients	Psychometric testing
SF-36	Deli et al, 2009[78]	-Limitations in physical activities because of health problems -Limitations in social activities because of physical or emotional problems -Limitations in usual role activities because of physical health problems -Bodily pain -General mental health -Limitations in usual role activities because of emotional problems -Vitality -General health perceptions	146	Correlation of the SF-36 vs the Axis II scales: Spearman coefficient. All items and subscales are significantly correlated with the exception of the JDC and the mental scales of the SF-36.
Social Support and Pain Questionnaire (SPQ)- Chinese	He and Wang, 2017[79]	–	118	Translation and cross-cultural adaptation Internal consistency: Cronbach α = 0.926 Test-retest reliability (2 wk): ICC = 0.784 Construct validity: EFA, 1 factor Convergent validity: Spearman rank correlation with global oral health question = 0.624 Significant correlation.
Social Support Scale	Funch et al, 1986[80]	-Perceptions of social support -Satisfaction with social support	92	Internal consistency: Cronbach α = 0.39 to 0.73 Criterion validity Construct validity: Correlation with CESD, POMS, TMAS
Symptom Severity Index-Modified (SSI)	Nixdorf et al, 2010[81]	-Jaw pain -Temple pain	108	Internal consistency: Cronbach α = 0.96 Dimensionality: EFA, 2 factors Test-retest reliability (55 patients, 2 to 48 h): ICC = 0.97 Between-item correlation: substantial but variable
Tampa Scale for Kinesiophobia (TSK-TMD)–from original Dutch to English	Visscher et al, 2010[82]	-Activity avoidance -Somatic focus	301	Cross-cultural adaptation Factor structure: EFA, 2 factors Test-retest reliability (4 wk, 58 patients): ICC = 0.73 Convergent validity: Pearson correlation with the Catastrophizing scale of the CSQ (Dutch version) = 0.23 Internal consistency: Cronbach α = 0.83.
TSK-TMD–Chinese	He et al, 2016[83]	–	160	Translation and cross-cultural adaptation Internal consistency: Cronbach α = 0.919 Test-retest reliability (30 patients, 2 wk): ICC = 0.797 Content validity: Interviews with patients and an expert panel Construct validity: EFA, 2 factors Convergent validity: Pearson correlation with global oral health question = 0.458–0.563
TSK-TMD–Brazilian Portuguese	Aguiar et al, 2017[84]	–	100	Cross-cultural validity Internal consistency: Cronbach α = 0.78 Test retest: ICC = 0.51–0.75. Structural validity: CFA, 2 factors Construct validity: Spearman rank correlation with PCS = 0.48, PHQ-8 = 0.38, MFIQ = 0.43 Convergent validity/discriminant validity: Average variance extracted

Table 2 Details of the Psychometric Properties of some PROMs in a TMD Population *(continued)*

PROM	Study	Domains/factors	No. of TMD patients	Psychometric testing
TMD Pain Screening Instrument Long Version (LV) Shot Version (SV)	Gonzalez et al, 2011[85]	–	504	Internal reliability: Cronbach α, LV = 0.93, SV = 0.87 Rasch analysis: Sensitivity = 99%, specificity = 97% EFA Temporal stability: ICC LV = 0.79, SV = 0.83
VAS score of the PSA (patient-specific activities)	Rollman et al, 2010[86]	–	132	Reproducibility: ICC = 0.72 Responsiveness: Sensitivity = 0.85%, specificity = 0.84%
WHO-5 Well-being Index	Ismail et al, 2018[87]	–	92	Internal consistency: Cronbach α = 0.883 Concurrent validity: Spearman correlation with OHIP-49, *r* = 0.705 Significant association.

AVE = average variance extracted; CESD = Centre for Epidemiologic Studies-Depression instrument; CFA = confirmatory factor analysis; CR = composite reliability; CVR = content validity ratio; CCC = concordance correlation coefficient; EFA = exploratory factor analysis; HIT-6 = Headache Impact Test; ICC = intraclass correlation coefficient; MPI = Multidimensional Pain Inventory; NPS = Numeric Pain Scale; NPV = negative predictive value; PI subscale = pain intensity subscale; POMS = Profile of Mood States; PPV = positive predictive value; PRCS = Pain-Related Control Scale; PSEQ = Pain Self-Efficacy Questionnaire; TMAS = Taylor Manifest Anxiety Scale.