CRANIO®: THE JOURNAL OF CRANIOMANDIBULAR & SLEEP PRACTICE
https://doi.org/10.1080/08869634.2025.2464227

Taylor & Francis
Taylor & Francis Group

TMJ

OPEN ACCESS Check for updates

Clinimetric analysis of the visual analogue scale and pain free mouth opening in patients with muscular temporomandibular disorder

Ian Young PT DSc, OCS, RMSK[a], James Dunning DPT, PhD, FAAOMPT[b,c], Firas Mourad PT, PhD[d,e], James Escaloni DPT, FAAOMPT, RMSK[b,f], Paul Bliton DPT, OCS, FAAOMPT[b,g] and César Fernández-de-Las-Peñas PT, PhD[h]

[a]Physical Therapy, Tybee Wellness & Osteopractic, Tybee Island, GA, USA; [b]Manipulative Therapy, American Academy of Manipulative Therapy Fellowship in Orthopaedic Manual Physical Therapy, Montgomery, AL, USA; [c]Physical Therapy & Acupuncture, Montgomery Osteopractic Physical Therapy & Acupuncture, Montgomery, AL, USA; [d]Department of Health, LUNEX University of applied sciences, Differdange, Luxembourg; [e]Physical Therapy, Luxembourg Health & Sport Sciences Research Institute, Differdange, Luxembourg; [f]Physical Therapy, Wellward Regenerative Medicine, Lexington, KY, USA; [g]Physical Therapy, William Middleton VA Hospital, Madison, WI, USA; [h]Department of Physical Therapy, Occupational Therapy, Rehabilitation and Physical Medicine, Universidad Rey Juan Carlos, Madrid, Spain

ABSTRACT

Objective: Perform a clinimetric analysis of the visual analogue scale (VAS) and active pain-free mouth opening (PFMO) in patients with muscular temporomandibular disorder (mTMD).

Methods: Reliability (intraclass correlation coefficient=ICC), construct validity, responsiveness (area under the curve=AUC), minimal detectable change (MDC), and minimal clinically important difference (MCID) values were calculated.

Results: The VAS-24hr (ICC=0.59), VAS-7day (ICC= 0.54), and PFMO (ICC=0.86) exhibited acceptable reliability. Both the VAS (AUC=0.96) and PFMO (AUC=0.87) exhibited a high level of responsiveness. The MCID was 15.5mm (VAS-24 and VAS-7day) and 3.5mm (PFMO) in the improved group; and 27.5mm (VAS-24), 21mm (VAS-7day), and 6.6mm (PFMO) in the much-improved group. The MDC was 9.6mm (VAS-24), 9.5mm (VAS-7day), and 6.1mm (PFMO). All outcomes demonstrated strong construct validity (Pearson's r; p<0.001)

Conclusions: All three outcome measures demonstrated acceptable clinimetric properties in patients with mTMD at the 3-month follow-up. The MCID lies outside measurement error in all outcomes in the much-improved group.

Keywords
Temperomandibular disorder; pain; reliability; validity; responsiveness; MCID

Introduction

Temporomandibular disorders (TMDs) are considered a group of conditions that may cause signs and symptoms, such as orofacial pain and dysfunction of musculoskeletal origin [1]. TMDs are often complex, affecting the temporomandibular joint and its related structures including the articular disc, joint capsule and muscles of mastication [2]. The overall prevalence of TMD is estimated to be 31% in adults/elderly and 11% in children/adolescents [3]. Conservative management of muscular TMD (mTMD) includes strategies that impact anatomic structures directly related to the specific etiology, including the upper cervical spine, joint capsule, articular disc and muscles of mastication, specifically the superior and inferior head of the lateral pterygoid [4]. The visual analogue scale (VAS) and measurements of mouth opening are commonly used outcome measures in patients with mTMD [5,6]. Unfortunately, there is a lack of large-scale high-quality evidence supporting the clinimetric properties of these routinely used outcomes in patients with mTMD. Although the VAS has acceptable clinimetric properties in common musculoskeletal [7–10] and non-musculoskeletal [11] conditions, currently there is no reliable data on the VAS in patients with mTMD. In regard to measures of mouth opening, only two small studies have reported reliability of maximal mouth opening (MMO), ranging from 0.85 to 0.96 [12–14], and only one study indicated good reliability (0.78) of active pain-free mouth opening (PFMO) [12]. Minimal detectable change (MDC) values for MMO range from 6 to 14 mm [14,15]. Lastly, just one smaller high-quality analysis ($n = 61$; 100% female) reported the meaningful clinically important difference (MCID) of the VAS (range: 0 to 19 mm) and MMO (range: 2.5 to 2.7 mm).[5] A recent randomized clinical trial (RCT) compared these outcomes (VAS & PFMO) in mTMD

CONTACT Ian Young tybeewellness@gmail.com Physical Therapy, Tybee Wellness & Osteopractic, 8 Horsepen Point Drive, Tybee Island 31328, GA, USA

patients ($n = 120$) receiving treatment with an interocclusal appliance, NSAIDs and joint mobilization to the temporomandibular joint (TMJ) *or* upper cervical spine manipulation and electrical dry needling to the muscular/periarticular regions associated with mTMD [16]. The purpose of this study was to address the outcome data from this previous trial [16] and comprehensively analyze the reliability, construct validity and responsiveness of the VAS and measures of PFMO in this large sample of patients treated with mTMD.

Materials and methods

A secondary clinimetric analysis of a prior multicenter randomized clinical trial [16] was performed on consecutive individuals ($n = 120$) with chronic (i.e., greater than 3 months) mTMD. The original trial was approved by the ethics committee at [anonymised] and was prospectively registered (ClinicalTrials.gov:[anonymised]). To be eligible, patients had to be at least 18 years old and meet the following criteria: (1) a clinical diagnosis of mTMD consistent with the Revised TMD group 1 Muscle Disorders Diagnostic Algorithm [17], (2) TMD symptoms for at least 3 months, and (3) an intensity of TMD symptoms of at least 30 mm on the VAS (0–100-mm) [18,19]. Patients were treated once or twice per week over a 4-week period with either an interocclusal appliance, NSAIDs (diclofenac), and non-thrust joint mobilization to the TMJ *or* upper cervical spine manipulation and electrical dry needling (EDN) [16]. Exclusion criteria and specific treatment procedures can be found in the original trial [16].

Outcome measures

The primary outcome was jaw pain intensity over the last 7 days, as measured by the VAS. Secondary outcomes included jaw pain intensity over the past 24 hr (VAS), active pain-free mouth opening (mm), and the Global Rating of Change (GROC). VAS ratings were analyzed at 3-months follow-up. The VAS consists of a 100 mm line, whereby the left side (starting from 0 mm) represents "no pain" and the right side (ending with 100 mm) represents "the worst pain imaginable." Patients were asked to make a mark on the line at the position that best represents their average pain intensity over the last 7 days. The VAS is an efficient, reliable and valid method of measuring subjective pain intensity in various patient populations, to include TMD [20–23]. The minimal clinically important difference (MCID) for the VAS has been shown to be 9–11 mm [24,25], and the minimal detectable change (MDC) for pain related to TMD is 10 mm to 14 mm in patients with TMD/juvenile idiopathic arthritis [15]. The MCID of the VAS in patients with TMD has been reported to range from 0 mm to 6 mm for moderately improved patients (GROC +1 to + 3) and 2.0 mm to 19 mm for largely improved patients (GROC = +4 to + 7) with TMD [5].

Maximum mouth opening (MMO) or active, pain-free mouth opening (PFMO) are commonly used outcomes to measure functional improvements in patients with TMD [18,26,27]. In the original trial [16] PFMO was used and measured as follows: the patient was asked to open their mouth as wide as possible without causing pain, from a supine position. At the end position, the distance between the upper and lower central incisors was measured in mm, and the average was taken over three attempts. Test–retest reliability of MMO (ICC = 0.85) and PFMO (ICC = 0.78) has been found within various diagnostic categories of TMD [12,13]. Test–retest reliability (Pearson's r) for MMO has also been reported to vary between 0.90 and 0.96 [14]. Prior studies have identified 6 mm to 9 mm as the smallest detectable change on MMO [14], while an MCID of 2.5 mm and 2.7 mm has been reported [5]. However, in patients specifically with mTMD, the MDC and MCID for PFMO remains to be established.

Patients also completed a 15-point GROC scale described by Jaeschke et al. [28] The scale ranges from −7 (a very great deal worse) to zero (about the same) to + 7 (a very great deal better). Scores of + 3 to + 5 are commonly used to identify "improved" versus "stable" (−2 to + 2) patients for psychometric/clinimetric analyses [29–31].

Data analysis

We categorized patients into three non-mutually exclusive groups at the 3-month follow-up based on their GROC scores: (a) those scoring from −2 to + 2 were considered clinically stable (minimal to no change); (b) those scoring ≥+3 (improved); and (c) those scoring ≥+5 (much improved). Patient variables for the improved (GROC ≥3 points) and stable (GROC −2 to + 2) groups were compared at baseline using independent t-tests for continuous data and chi-square tests for categorical data. Patients could be classified into more than one group, as these different groups were used for one or more analyses of reliability, validity, and responsiveness. Our main analysis focused on patients who were stable and those who reported being improved at the 3-month follow-up, whereas the groups reporting being much improved were used for clinical comparative analysis.

Test–retest reliability was examined for the VAS-24 hr, VAS-7day, and PFMO using "stable" patients by comparing scores at the initial examination with those at the 3-month follow-up. The intra-class correlation coefficient (ICC) was calculated according to the procedure described by Shrout and Fleiss [32]. Values <0.50 indicate poor reliability, while values between 0.50–0.75, 0.75–0.90, and >0.90 denote moderate, good, and excellent agreement, respectively.

Construct validity of the VAS-24 hr, VAS-7day, and PFMO was examined by comparing the change in outcome scores for the "'stable'" and "improved" groups using separate, two-way mixed-model analyses of variance for the repeated measures at baseline and reevaluation. We hypothesized that "'stable'" patients in each group would have VAS-24 hr, VAS-7day, and PFMO intake values that did not change, whereas patients classified in the improved categories would demonstrate a significant change in their values. This would be represented by a significant group x time interaction. Pearson's *r* correlation coefficient was also examined between all outcome measures.

Responsiveness, the ability of a measure to recognize change when change has occurred, was assessed for the VAS-24 hr, VAS-7day, and PFMO using the clinically "'stable,'" and "improved" groups at the 3-month follow-up point. Receiver operator characteristic (ROC) curves [33] were constructed by plotting sensitivity values (true-positive rate) on the y-axis and 1-specificity values (false-positive rate) on the x-axis for each level of change score. Separate ROC curves were constructed for the VAS-24 hr, VAS-7day, and PFMO. The area under the curve (AUC) and the 95% CI were obtained as a method for determining the ability of each measure to distinguish improved patients from stable patients in each category. An AUC of 0.50 indicates that the measure has no diagnostic accuracy beyond chance, whereas a value of 1 suggests perfect accuracy [33]. MCID, the smallest difference that patients perceive as beneficial, was calculated by identifying the point on the ROC curve nearest to the upper left-hand corner, which is considered to be the best cutoff score for distinguishing improved and stable patients [22]. Sensitivity and specificity values for the selected cutoff scores were also calculated. MDC, the amount of change that must be observed before the change can be considered to have exceeded measurement error, was calculated by determining the standard error of measurement (SEM) for the VAS-24 hr, VAS-7day, and PFMO in the stable group (n = 23) from the ICC reliability analysis. The SEM was calculated with the formula SD/∏n, where SD is the standard deviation of the change score values and *n*= the sample size. The SEM was multiplied by 1.65 to determine the 90% CI (MDC90), and then multiplied by the ∏2 to account for the errors taken with repeated measurements.

Results

One hundred and twenty patients satisfied the inclusion and exclusion criteria, completed the study, and were included in data analysis. Baseline characteristics are located in Table 1. The mean GROC score for all patients included in the analysis at the 3-month follow-up was + 3.8 (SD + 2.6). The mean GROC score for the improved vs. stable groups was + 5.2 (SD + 1.4) and + 0.9 (SD + 1.2), respectively. At the 3-month follow-up, 83 (69.2%) patients were classified as improved, and 35 (29.2%) remained stable. There was a significant difference ($p < .001$) in mean change scores between stable and improved patients for the VAS-24 hr, VAS-7day, and PFMO, at the 3-month follow-up (Table 2). Additionally, all three outcome measures exhibited strong construct validity (Pearson's r ranging from 0.57 to 0.85, Table 4)

The test–retest (ICC) values and MDC calculated from the stable patients are reported in Table 3. At the 3-month follow-up, the VAS-24 hr (ICC: 0.59, 95%CI 0.20–0.80) and VAS-7 day (ICC: 0.54, 95%CI 0.16–0.76) had fair reliability, while PFMO had excellent reliability (ICC: 0.86, 95%CI 0.67–0.94). At 3-months, the MDC was 9.6 mm, 9.5 mm and 6.1 mm for the VAS-24 hr, VAS-7 day and PFMO, respectively.

The VAS-24 hr, VAS-7 day and PFMO all demonstrated excellent responsiveness (AUC range: 0.87 to 0.96) and are reported in Table 3. Individual ROC curves can also be found in Figure 1. The MCID threshold and the sensitivity/specificity associated with each cutoff score are also located in Table 3. In the improved group (GROC +3 to + 7), the MCID was 15.5 mm for the VAS-24 hr, 15.5 mm for the VAS-7 day and 3.5 mm for PFMO. In the much-improved group (GROC +5 to + 7), the MCID was 27.5 mm for the VAS-24 hr,

Table 1. Baseline characteristics.

Outcome Measures	*N* = 120
Gender: male/female	30/90
Age: yrs.	41.6 ± 12.8
Weight: kg	73.0 ± 14.5
Years with jaw pain	7.1 ± 7.9
Visual Analogue Scale – 24hrs (0–100)	49.2 ± 14.3
Visual Analogue Scale – 7 days (0–100)	53.7 ± 13.6
Active Pain Free Mouth Opening (mm)	32.1 ± 7.3

Table 2. Difference between change scores from baseline to 3-months follow-up on outcomes measures.

Outcome Measure	Improved GROC (+3 to + 7) *N* = 83 Mean (SD)	Stable GROC (−2 to + 2) *N* = 35 Mean (SD)	Mean Difference (95% CI)	*P*
VAS-Past 24hr (mm)	35.43 (16.8)	7.37 (10.7)	28.1 (22; 34.2)	$p < .001$
VAS-Past 7 days (mm)	39.06 (18.0)	7.91 (11.5)	31.2 (25.1; 37.2)	$p < .001$
Active Pain Free Mouth Opening (mm)	10.24 (7.1)	2.06 (3.6)	8.9 (2.4; 14.0)	$p < .001$

VAS = visual analogue scale (0–100), Active Pain Free Mouth Opening (mm), GROC = global rating of change, SD=standard deviation, CI = confidence interval.

Table 3. Clinimetric properties of patient-rated outcome measures used for temporomandibular disorder.

Outcome Measure	AUC 95% CI Improved (GROC +3 to + 7) *N* = 83	MCID Sn; Sp Improved (GROC +3 to + 7) *N* = 83	AUC 95% CI Much Improved (GROC +5 to + 7) *N* = 60	MCID Sn; Sp Much Improved (GROC +5 to + 7) *N* = 60	ICC 95% CI Stable (GROC= −2 to + 2) *N* = 35	MDC
VAS-24hr (mm)	0.93 0.88; 0.98	15.5 0.94; 0.84	0.96 0.92; 0.99	27.5 0.90; 0.90	0.59 0.20, 0.80	9.6
VAS-7 days (mm)	0.94 0.90; 0.99	15.5 0.95; 0.81	0.96 0.93; 0.99	21.0 0.95; 0.87	0.54 0.16; 0.76	9.5
Active Pain Free Mouth Opening (mm)	0.86 0.79; 0.92	3.5 0.81; 0.73	0.87 0.80; 0.94	6.5 0.83; 0.82	0.86 0.67; 0.94	6.1

VAS = visual analogue scale (0–100), AUC = area under the curve, GROC = global rating of change, MCID = minimally clinically important difference, Sn = sensitivity, Sp = specificity, ICC = intraclass correlation coefficient, CI = confidence interval, MDC_{90} = minimal detectable change.

Table 4. Pearson's correlation coefficient (r).

Outcome Measures	VAS-24hr s r (95% CI)	VAS-7 day s r (95% CI)	GROC r (95% CI)
Pain Free Mouth Opening	0.57 (0.44; 0.68) *p <.001*	0.60 (0.48; 0.71) *p <.001*	0.61 (0.48; 0.71) *p <.001*
VAS-24 hr	—–	0.85 (0.79; 0.86) *p <.001*	0.76 (0.68; 0.82) *p <.001*
VAS-7 days	—–	—–	0.82 (0.75; 0.87) *p <.001*

VAS = visual analogue scale (0–100), GROC = global rating of change, GROC = global rating of change scale, CI-confidence interval.

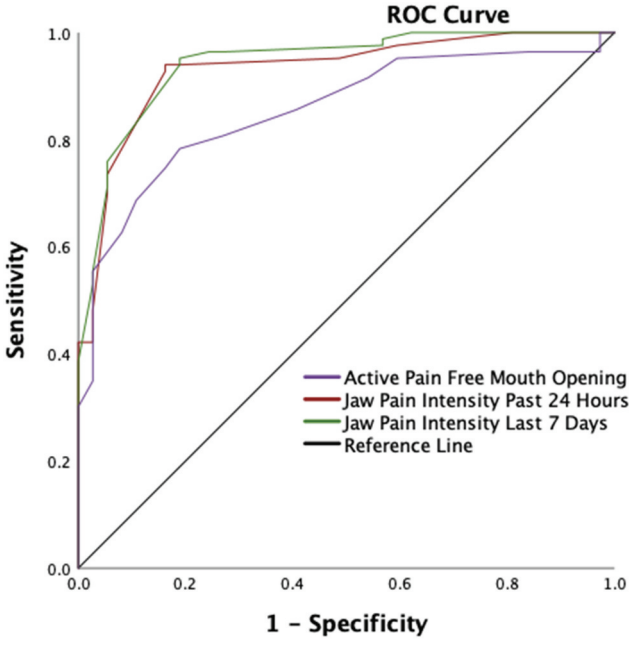

Figure 1. Receiver operating curves for all outcomes.

21.0 mm for the VAS-7 day and 6.5 mm improvement for PFMO.

Discussion

This study comprehensively examined the clinimetric properties of the VAS (24 hr and 7day) and PFMO in patients successfully treated with chronic mTMD at a 3-month follow-up. The VAS-24 hr, VAS-7 day and PFMO exhibited acceptable reliability, strong construct validity, and a high-level of responsiveness in this patient population (Tables 2–4, Figure 1).

Regarding reliability of the VAS, no prior studies have reported data for patients with chronic mTMD. Our results found only moderate reliability of the VAS-24 hr (ICC = 0.59) and VAS-7 day (0.54) in this patient population, and this may be a result of self-report dependency [34]. Although the moderate reliability of the VAS may be statistically acceptable, clinicians and researchers should use caution when solely using PROMs in patients with TMDs. Using both PROMs

and physical objective outcomes including MMO/PFMO may give a better estimate of improvement. Prior test–retest, reliability of MMO (ICC = 0.85) and PFMO (ICC = 0.78) has been reported within mixed diagnostic categories of TMD [12], while a single dated study examined 25 patients with painfully restricted TMD and reported reliability (Pearson's *r*) to be between 0.90 and 0.96 [14]. Unfortunately, Pearsons' *r* only gives a sense of "relative reliability", lacking clinical discernment to systematic measurement errors [35]. Our data suggests that measures of PFMO have excellent reliability (ICC = 0.86) at the 3-month follow-up in patients with mTMD (Table 3).

In the current study, the responsiveness values of the VAS (AUC range: 0.93 to 0.94) in the improved group (GROC: +3 to + 7) are in contrast with prior clinimetric analysis of a moderately improved group (GROC: +1 to + 3) with TMD (VAS: AUC range 0.62 to 0.69) [5]. However, in the same analysis [5], the largely improved group (GROC +4 to + 7) had more comparable responsiveness of the VAS (AUC range = 0.81 to 0.87) to the much improved group (GROC:+5 to + 7) in the current analysis (AUC = 0.96). The responsiveness of mouth opening measures in the current study (PFMO: AUC = 0.86; Table 2) is also in contrast to prior findings (MMO: AUC range = 0.57 to 0.65) [5]. To compare and contrast, the Calixtre et al [5] study included *n* = 61 female patients (18 = moderately improved; 13 = largely improved), while the current study included a larger sample of 120 male and female patients (83 = improved; 63 = much improved: Tables 1 and 2). Furthermore, the analysis by Calixtre et al [5] was performed based on a 5-week follow-up rather than the 3-month follow-up used in the current study.

In the current study, the MDC/MCID of the VAS-24 hr was 9.6 mm/15.5 mm in the improved group and 9.6 mm/27.5 mm in the much-improved group. The MDC/MCID for the VAS-7 day was 9.5 mm/15.5 mm in the improved group and 9.5/27.5 in the much-improved group. The smallest detectable change (10 mm to 14 mm) has been assessed in 33 TMD patients with juvenile idiopathic arthritis [15]. Importantly, the aforementioned study had a mean age of 14 year [15] and may not be generalizable with the findings of our current study (mean age = 42 years). Notably, the Calixtre et al [5] study did not report the MDC; however, the MCID of the VAS ranged from 0 mm to 6 mm for the moderately improved group and 2.0 mm to 19 mm in the largely improved group. Regarding mouth opening measures, the MDC/MCID of PFMO in the current study was 6.1 mm/3.5 mm in the improved group and 6.1 mm/6.5 mm in the much-improved group. Kropmans et al [14] reported 6–9 mm as the "smallest detectable difference" of MMO in 25 patients with TMD at the one week follow-up, while Calixtre et al [5] reported the MCID of MMO was 2.5 mm for the moderately improved group and 2.7 mm in the largely improved group. The contrast in sample size, age/gender distribution, statistical methods/description of the MDC, and follow-up time frames may be components swaying these differences between studies. Furthermore, the inclusion of "pain-free" mouth opening in the current study vs. "maximal" mouth opening in prior studies may have had an impact on the difference in measurements, and hence the clinimetric outcomes. We suggest readers to refer to Table 3 for a summary of the clinimetric properties of VAS and PFMO in patients with TMD.

In regard to construct validity, all three outcome measures demonstrated strong correlation with each other (Table 4). Additionally, the original RCT, Dunning et al [16] found significant between group differences in pain, PFMO, and the GROC, favoring the spinal manipulation and EDN group. In contrast, a small RCT [36] investigated the addition of cervical spine thrust-manipulation to a treatment program of behavioral education, soft tissue mobilization and home exercise in patients with TMD. The results suggested there were significant differences between groups in PROMs of jaw function, fear of movement, and the GROC favoring the spinal manipulation group [36]. However, there was no interaction between groups with MMO and the NPRS [36]. Although the RCT by Reynolds et al [36] lacks clinimetric analysis, and limited statistical comparisons can be made to the current study, the differences in sample size (*n*-120[16] vs. n = 50[36]), baseline jaw pain scores (5.4[16] vs. 3.7[36]), the use of PFMO [16] vs. MMO [36], a more robust treatment group treating pain mechanisms (spinal manipulation + EDN) [16], and follow-up time frames (3-month [16] vs. 4-week [36] may have had an effect on the difference in outcomes between the two RCTs [16,36], and positive clinimetric outcomes of the current study.

Study limitations

The current analysis does not come without limitations. First, our results may not be generalizable to all gender sample imbalances, subgroups, diagnostic categories, and available interventions/outcome measures in patients with mTMD. Our analysis included only those patients with mTMD involved in the original randomized clinical trial, including two standardized treatment interventions and two standardized outcome measures. Second, the results are short-term in nature, not expanding beyond 3-month follow-up. Third, self-reported pain scores could introduce bias due to

subjective interpretation. Future research should address the above limitations, and any innovative tools used for assessing outcomes in patients with mTMD.

Conclusion

To date, this is the largest comprehensive clinimetric analysis of pain (VAS) and PFMO in patients treated with mTMD. Our results suggest that the VAS and PFMO both exhibited acceptable reliability, strong construct validity and a high level of responsiveness over time. To be considered clinically meaningful, clinicians and researchers should expect a 15.5 mm change on both VAS scales, and a 3.5 mm increase with PFMO in improved patients at the 3-month follow-up. In patients that reported being much-improved, a 21 mm change (VAS-24 hr), a 27.5 mm change (VAS-7 day), and a 6.5 mm improvement with PFMO should be expected. The MCID found in improved group should be considered, as it lies outside measurement error in all three outcomes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Ian Young PT DSc, OCS, RMSK http://orcid.org/0000-0001-6559-993X

References

[1] Manfredini D, Häggman-Henrikson B, Al Jaghsi A, et al. Temporomandibular disorders: INfORM/IADR key points for good clinical practice based on standard of care. CRANIO®. 1–5. doi: 10.1080/08869634.2024.2405298

[2] Butts R, Dunning J, Perreault T, et al. Pathoanatomical characteristics of temporomandibular dysfunction: Where do we stand? (Narrative review part 1). J Bodyw Mov Ther. 2017 Jul;21(3):534–540. doi: 10.1016/j.jbmt.2017.05.017

[3] Valesan LF, Da-Cas CD, Réus JC, et al. Prevalence of temporomandibular joint disorders: a systematic review and meta-analysis. Clin Oral Investig. 2021 Feb;25(2):441–453. doi: 10.1007/s00784-020-03710-w

[4] Butts R, Dunning J, Pavkovich R, et al. Conservative management of temporomandibular dysfunction: A literature review with implications for clinical practice guidelines (Narrative review part 2). J Bodyw Mov Ther. 2017 Jul;21(3):541–548. doi: 10.1016/j.jbmt.2017.05.021

[5] Calixtre LB, Oliveira AB, Alburquerque-Sendín F, et al. What is the minimal important difference of pain intensity, mandibular function, and headache impact in patients with temporomandibular disorders? Clinical significance analysis of a randomized controlled trial. Musculoskelet Sci Pract. 2020 Apr;46:102108. doi: 10.1016/j.msksp.2020.102108

[6] Taimeh D, Ni Riordain R, Fedele S, et al. Patient-reported outcome measures used in temporomandibular disorders: A review of the literature. J Oral Facial Pain Headache. 2023;37(2):113–129. doi: 10.11607/ofph.3264

[7] Chiarotto A, Maxwell LJ, Ostelo RW, et al. Measurement properties of visual analogue scale, numeric rating scale, and pain severity subscale of the brief pain inventory in patients with low back pain: a systematic review. J Pain. 2019 Mar;20(3):245–263. doi: 10.1016/j.jpain.2018.07.009

[8] Modarresi S, Lukacs MJ, Ghodrati M, et al. A systematic review and synthesis of psychometric properties of the numeric pain rating scale and the visual analog scale for use in people with neck pain. Clin J Pain. 2021 Oct 26;38(2):132–148. doi: 10.1097/ajp.0000000000000999

[9] Alghadir AH, Anwer S, Iqbal A, et al. Test-retest reliability, validity, and minimum detectable change of visual analog, numerical rating, and verbal rating scales for measurement of osteoarthritic knee pain. J Pain Res. 2018;11:851–856. doi: 10.2147/jpr.S158847

[10] Beletsky A, Nwachukwu BU, Gorodischer T, et al. Psychometric properties of visual analog scale assessments for function, pain, and strength compared with disease-specific upper extremity outcome measures in rotator cuff repair. JSES Int. 2020;4(3):619–624. doi: 10.1016/j.jseint.2020.04.012

[11] Karcioglu O, Topacoglu H, Dikme O, et al. A systematic review of the pain scales in adults: Which to use? Am J Emerg Med. 2018 Apr;36(4):707–714. doi: 10.1016/j.ajem.2018.01.008

[12] Goulet JP, Clark GT, Flack VF, et al. The reproducibility of muscle and joint tenderness detection methods and maximum mandibular movement measurement for the temporomandibular system. J Orofac Pain. 1998;12(1):17–26.

[13] Rauch A, Schierz O. Reliability of mandibular movement assessments depending on TMD. Cranio. 2018 May;36(3):156–160. doi: 10.1080/08869634.2017.1325986

[14] Kropmans T, Dijkstra P, Stegenga B, et al. Smallest detectable difference of maximal mouth opening in patients with painfully restricted temporomandibular joint function. Eur J Oral Sci. 2000 Feb;108(1):9–13. doi: 10.1034/j.1600-0722.2000.00747.x

[15] Stoustrup P, Kristensen KD, Verna C, et al. Orofacial symptoms related to temporomandibular joint arthritis in juvenile idiopathic arthritis: smallest detectable difference in self-reported pain intensity. J Rheumatol. 2012 Dec;39(12):2352–2358. doi: 10.3899/jrheum.120437

[16] Dunning J, Butts R, Bliton P, et al. Dry needling and upper cervical spinal manipulation in patients with

temporomandibular disorder: A multi-center randomized clinical trial. Cranio. 2022 Apr 12:1–14. doi: 10.1080/08869634.2022.2062137

[17] Look JO, Schiffman EL, Truelove EL, et al. Reliability and validity of axis I of the research diagnostic criteria for temporomandibular disorders (RDC/TMD) with proposed revisions. J Oral Rehabil. 2010 Oct;37 (10):744–759. doi: 10.1111/j.1365-2842.2010.02121.x

[18] R LT, Fernández-DE-Las-Peñas C, Fernández-Carnero K, et al. The effects of manual therapy and exercise directed at the cervical spine on pain and pressure pain sensitivity in patients with myofascial temporomandibular disorders. J Of Oral Rehabil. 2009;36 (9):644–652. doi: 10.1111/j.1365-2842.2009.01980.x

[19] Gonzalez-Iglesias J, Cleland JA, Neto F, et al. Mobilization with movement, thoracic spine manipulation, and dry needling for the management of temporomandibular disorder: a prospective case series. Physiother Theory Pract. 2013 Nov;29(8):586–595. doi: 10.3109/09593985.2013.783895

[20] Bijur PE, Silver W, Gallagher EJ. Reliability of the visual analog scale for measurement of acute pain. Acad Emerg Med. 2001 Dec;8(12):1153–1157. doi: 10.1111/j.1553-2712.2001.tb01132.x

[21] Le Resche L, Burgess J, Dworkin SF. Reliability of visual analog and verbal descriptor scales for "objective" measurement of temporomandibular disorder pain. J Dent Res. 1988 Jan;67(1):33–36. doi: 10.1177/00220345880670010601

[22] Conti PC, de Azevedo Lr, de Souza Nv, et al. Pain measurement in TMD patients: evaluation of precision and sensitivity of different scales. J Oral Rehabil. 2001 Jun;28 (6):534–539. doi: 10.1046/j.1365-2842.2001.00727.x

[23] Stuginski-Barbosa J, Silva R, Cunha C, et al. Pressure pain threshold and pain perception in temporomandibular disorder patients: is there any correlation? Rev Dor. 2015;16(1). doi: 10.5935/1806-0013.20150005

[24] Bird SB, Dickson EW. Clinically significant changes in pain along the visual analog scale. Ann Emerg Med. 2001 Dec;38(6):639–643. doi: 10.1067/mem.2001.118012

[25] Gallagher EJ, Liebman M, Bijur PE. Prospective validation of clinically important changes in pain severity measured on a visual analog scale. Ann Emerg Med. 2001 Dec;38(6):633–638. doi: 10.1067/mem.2001.118863

[26] Naeije M, Te Veldhuis AH, Te Veldhuis EC, et al. Disc displacement within the human temporomandibular joint: a systematic review of a 'noisy annoyance'. J Oral Rehabil. 2013 Feb;40(2):139–158. doi: 10.1111/joor.12016

[27] Fernandez-Carnero J, La Touche R, Ortega-Santiago R, et al. Short-term effects of dry needling of active myofascial trigger points in the masseter muscle in patients with temporomandibular disorders. J Orofac Pain. 2010 Winter;24(1):106–112.

[28] Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989 Dec;10 (4):407–415. doi: 10.1016/0197-2456(89)90005-6

[29] Young IA, Cleland JA, Michener LA, et al. Construct validity, and responsiveness of the neck disability index, patient-specific functional scale, and numeric pain rating scale in patients with cervical radiculopathy. Am J Phys Med & Rehabil. 2010;89(10):831–839. doi: 10.1097/PHM.0b013e3181ec98e6

[30] Young Ia D, Dunning J, Butts R, et al. Reliability, construct validity, and responsiveness of the neck disability index and numeric pain rating scale in patients with mechanical neck pain without upper extremity symptoms. Physiother Theory Pract. 2018 Jun 1;35 (12):1–8. doi: 10.1080/09593985.2018.1471763

[31] Young IA, Dunning J, Butts R, et al. Psychometric properties of the numeric pain rating scale and neck disability index in patients with cervicogenic headache. Cephalalgia. 2019 Jan;39(1):44–51. doi: 10.1177/0333102418772584

[32] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979 Mar;86 (2):420–428. doi: 10.1037//0033-2909.86.2.420

[33] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982 Apr;143(1):29–36. doi: 10.1148/radiology.143.1.7063747

[34] Chang EM, Gillespie EF, Shaverdian N. Truthfulness in patient-reported outcomes: factors affecting patients' responses and impact on data quality. Patient Relat Outcome Meas. 2019;10:171–186. doi: 10.2147/prom.S178344

[35] Vaz S, Falkmer T, Passmore AE, et al. The case for using the repeatability coefficient when calculating test–retest reliability. PLUS ONE. 2013;8(9):e73990. doi: 10.1371/journal.pone.0073990

[36] Reynolds B, Puentedura EJ, Kolber MJ, et al. Effectiveness Of cervical spine high-velocity, low-amplitude thrust added to behavioral education, soft tissue mobilization, and exercise for people with temporomandibular disorder with myalgia: A randomized clinical trial. J Orthop Sports Phys Ther. 2020 Aug;50(8):455–465. doi: 10.2519/jospt.2020.9175