

Validity and reproducibility of the TMD Assessment Test: diagnostic accuracy study

Validade e reprodutibilidade do Teste Avaliativo de DTM: um estudo de acurácia diagnóstica

Yasmin Bianca Oliveira Neri¹ Ana Carolina Silva Francisco² Isabela Raul Nepomuceno de Jesus³ Italo Eduardo Novaes Santos⁴ Rafael Santos Ferreira⁵ Ana Lúcia Barbosa Góes⁶ ¹Corresponding author. Universidade Federal da Bahia (Salvador). Bahia, Brazil. yasmineri6@gmail.com²⁻⁵Universidade Federal da Bahia (Salvador). Bahia, Brazil. anacarollinafisio@outlook.com, isabelaraulndj@gmail.com, italoeduardo03@yahoo.com
rsferreirars@hotmail.com⁶Universidade Federal da Bahia (Salvador); Escola Bahiana de Medicina e Saúde Pública (Salvador). Bahia, Brazil. goes.ana@ufba.br

ABSTRACT | INTRODUCTION: Temporomandibular disorders are TMJ disorders of muscular, articular, or mixed origin and closely related to postural alterations. Diagnostic tools have gaps regarding clinical application and do not associate posture. **OBJECTIVE:** Validate the TMD Assessment Test (TMDAsT) regarding its diagnostic accuracy and reproducibility. **MATERIALS AND METHODS:** Diagnostic accuracy study with individuals between 18 and 59 years old evaluated by the Fonseca Assessment Index (FAI), sociodemographic questionnaire, and TMDAsT, the latter performed by three different trained examiners. The result of FAI diagnosis categorization was compared with the result of TMDAsT. A Chi-square test was used for diagnostic accuracy. Positive predictive values (PPV) and negative predictive values (NPV) were determined. The Kappa of Fleiss did the reproducibility between three examiners. Cohen's Kappa, for 2x2 analysis. All tests with 5% significance. **RESULTS:** Of the 10 participants assessed, FAI identified 8 participants with a TMD diagnosis while the TMDAsT verified 9 participants with this dysfunction. Sensitivity was 100%, specificity 50%, PPV 88% and NPV 50%. Fleiss' Kappa showed reasonable reliability (K = 0.26 [95% CI: -0.099 - 0.617]; p>0.05). Cohen's Kappa showed insignificant reproducibility between observers 1 and 2 (K=-0.11; p>0.05; discordance=80%), and 1 and 3 (K= -0.11; p>0.05; discordance=80%), perfect reproducibility between observers 2 and 3 (K=1.00; p<0.05; concordance=100%). **CONCLUSION:** TMDAsT presents high sensitivity and low specificity but with low reproducibility until the present moment.

KEYWORDS: Temporomandibular joint disorders. Survey and questionnaires. Diagnostic accuracy. Reproducibility of results.

RESUMO | INTRODUÇÃO: Disfunções temporomandibulares (DTM) são distúrbios na ATM, sendo de origem muscular, articular ou mista, e com íntima relação com alterações posturais. Instrumentos de diagnóstico apresentam lacunas quanto à aplicação clínica e não associam à postura. **OBJETIVO:** Validar o Teste Avaliativo de DTM (TAvDTM) quanto à acurácia diagnóstica e reprodutibilidade. **MATERIAIS E MÉTODOS:** Estudo de acurácia diagnóstica com indivíduos entre 18 e 59 anos avaliados pelo Índice Anamnésico de Fonseca (IAF), questionário sociodemográfico e TAvDTM, este último realizado por três examinadores diferentes e treinados. O resultado da categorização do diagnóstico do IAF foi comparado com o resultado do TAvDTM. Para acurácia diagnóstica utilizou-se teste Qui-Quadrado. Valores Preditivos Positivos (VPP) e Negativos (VPN) foram determinados. A reprodutibilidade entre os três examinadores foi feita por meio do Kappa de Cohen, para análise 2x2 e de Fleiss. Todos os testes com significância de 5%. **RESULTADOS:** Dos 10 participantes avaliados, o IAF identificou 8 com diagnóstico de DTM, enquanto o TAvDTM verificou 9 participantes com presença desta disfunção. A sensibilidade foi de 100%, especificidade de 50%, VPP de 88% e VPN de 50%. O Kappa de Fleiss evidenciou confiabilidade razoável (K = 0,26 [IC 95%: -0,099 - 0,617]; p>0,05). O Kappa de Cohen mostrou reprodutibilidade insignificante entre os avaliadores 1 e 2 (K=-0,11; p>0,05; discordância=80%), e 1 e 3 (K= -0,11; p>0,05; discordância=80%), reprodutibilidade perfeita entre os avaliadores 2 e 3 (K=1,00; p<0,05; concordância=100%). **CONCLUSÃO:** O TAvDTM apresenta alta sensibilidade e baixa especificidade, porém com baixa capacidade de reprodução até o presente momento.

PALAVRAS-CHAVE: Transtornos da Articulação Temporomandibular. Inquéritos e questionários. Acurácia diagnóstica. Reprodutibilidade dos testes.

How to cite this article: Neri YBO, Francisco ACS, Jesus IRN, Santos IEN, Ferreira RS, Góes ALB. Validity and reproducibility of the TMD Assessment Test: diagnostic accuracy study. J Physiother Res. 2021;11(4):774-782. <http://dx.doi.org/10.17267/2238-2704rpf.v11i4.4153>

Introduction

Temporomandibular dysfunctions (TMD) are defined as a set of joint and/or muscle disorders that lead to difficulties in the functioning of the temporomandibular joint (TMJ). It has a multifactorial etiology and is associated with trauma (mandibular and TMJ), parafunctional habits (chewing gum, biting lips and cheeks, gnawing objects, clenching or grinding teeth), postural deviations, among others.¹

TMD is often diagnosed in the population aged between 19 and 40 years.² This disorder has an incidence in the world population, in which 65% manifest at least one symptom and 35% a sign of this dysfunction³, with prevalence in females⁴ and a higher probability of presenting symptoms such as pain in the neck and shoulders, facial muscles, temporomandibular joint and headache.²

As for posture, the severity of these changes is related to the flexibility of the posterior muscle chain.² Postural deviations in the cervical spine, head, shoulders, and other segments may be causal factors of craniocervical dysfunction, subsequently causing the perpetuation of TMD symptoms and signs.^{2,6,7} It was also identified that the muscles responsible for mastication are related to body posture through complex neuromuscular connections. Thus, it is understood that there is a relationship between TMJ changes and postural alignment.^{2,7}

Considering posture as a complex ability based on the interaction of dynamic sensory-motor processes and understanding that the relationship of bone parts depends on the interaction of various muscle chains⁸, it is possible to consider the existence of the association between TMD and muscle chains, which are muscles that work in synergy for the maintenance of posture and proper functioning of body structures. Thus, if one area is compromised, it can affect other regions related to these muscle chains.

In the literature, there are several instruments⁹⁻¹² commonly adopted in the field of physiotherapy and dentistry in order to obtain a correct diagnosis

of TMD and thus offer the most appropriate treatment. The Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) is one of the most widely used questionnaires in clinical and research settings since its publication in 1992. However, like other instruments, this questionnaire has its shortcomings, such as difficulties in clinical practice because it is a research-oriented questionnaire that is very complex and has a long application time.¹¹

As an example of easy-to-administer tools, we have the Diagnostic Criteria for Temporomandibular Disorders (DC/TMD), which is used for both research and clinical practice and is currently considered the gold standard, and the Fonseca Anamnesis Index (FAIx), which is used to determine the severity of TMD according to its signs and symptoms and has been widely used as a screening tool since its validation. However, these indices do not consider posture as a variable to be included in the evaluation.¹³

The growing relationship between posture and temporomandibular dysfunction as presented in literature⁶ is notorious. Thus, the Temporomandibular Dysfunction Evaluation Test was proposed, consisting of five postural movements (trunk flexion, right lateral tilt, left lateral tilt, right lateral torsion, and left lateral torsion), in which the evaluator performs these movements passively and analyzing the presence of muscle limitations. This test considers the posterior muscle chain tension analysis, which aims to identify individuals with or without temporomandibular dysfunction.

Since this is a test that has not been validated in the literature, its validation is necessary before implementing it in clinical practice or research environments. Thus, this study aimed to validate the Temporomandibular Dysfunction Evaluation Test in terms of its diagnostic accuracy and to provide health professionals with a validated tool that, in addition to complementing the diagnosis of TMD, allows the relationship with postural changes, enabling health professionals to offer adequate treatment to the population.

Methods

A prospective, blind, diagnostic accuracy study was conducted from October 2019 to March 2020. All research participants signed the Informed Consent Form (ICF), according to resolution 466/12 of the National Health Council. CEP ICS-UFBA approved this research under CAAE nº26305519.4.0000.5662.

This study was conducted to identify the agreement between the research examiners to apply the Temporomandibular Dysfunction Evaluation Test better since it is an evaluator-dependent instrument.

The inclusion criteria used were individuals of both genders aged between 18 and 59 years. As exclusion criteria, individuals with stroke or who had undergone oral and maxillary surgery, or individuals who had facial trauma, presence of rheumatoid arthritis, diagnosis of scoliosis, hip and spine degeneration, who had undergone physiotherapy or orthodontic treatment for temporomandibular dysfunction, who presented tooth loss or were taking medication (analgesics, anti-inflammatory, and muscle relaxants).

Initially, an evaluator applied a sociodemographic questionnaire, and then the FAIx was used as the gold standard in comparing the tests. This index has ten items with possible answers: yes (10 points), sometimes (5 points), and no (0 points), and its classification is defined as no TMD (0 to 15 points), mild TMD (20 to 45 points), moderate TMD (50 to 65) and severe TMD (70 to 100 points). It was not possible to use the CD/TMD, the current gold standard for TMD diagnosis, because it was still being validated during the collection period of this study.

After administering the tests, the participants were analyzed with the TMD Assessment Test, in random order, by three properly trained examiners, and the evaluations occurred with a 15-minute interval between examiners to avoid muscle stimulation and not cause interference in the results.

According to Busquet's philosophy and theory of muscle chains, posterior chain movements indicate a relationship between posture and TMD, while the anterior muscle chain indicates a relationship with organ dysfunction or visual dysfunction.

To carry out the TMD Evaluation Test, the research included one evaluator, who controlled the examiners; three examiners applied the movement patterns that cause tension in the posterior muscle chain and determined the presence or absence of TMD. The possible diagnosis for the instrument would be negative, one cross (+) or two crosses (++). Those who received from (++) in any movement pattern were diagnosed with temporomandibular dysfunction, and individuals who received (+) or negative were considered without the presence of TMD, being (+) indicative of mild muscle resistance to the movements performed, not interfering in the diagnosis.

Importantly, the participants were identified from their registration number, ensuring anonymity, and the examiners were blinded from each other, as each was in an isolated room, with no contact from one examiner to the other.

The instructions for the test were: 1- The individual should be in orthostatic position, with feet together, barefoot, with a relaxed posture and looking straight ahead; 2- The movements were performed only by the examiner, that is, passively, and if the individual tried to help the movement, it would be interrupted, and the movement would start from the initial posture. The movement was repeated as many times as necessary for diagnosis; 3- If the individual wore glasses, he/she should keep wearing them during the test.

To perform the TMD Evaluation Test (figure 1), the examiner observed the posterior muscle tensions produced by the following movements:

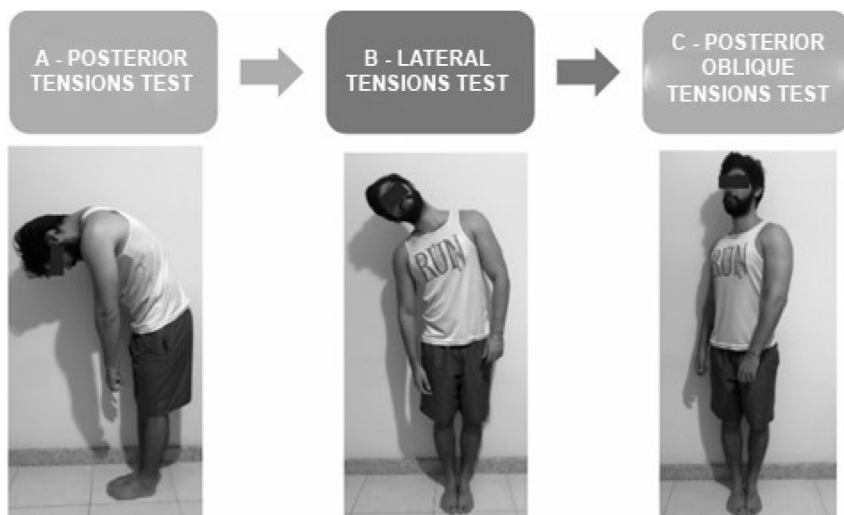
A- Flexion of the trunk or posterior tensions test - promoting rolling of the participant's entire body starting from the head and progressing slowly and smoothly throughout the spine.

B- Right and left lateral trunk tilts or lateral tensions test- allows for lateral tilting of the participant's entire body starting from the head and progressing slowly and smoothly down the entire spine.

C- Right anterior twist and left anterior twist of the trunk or posterior oblique tension test - promotes anterior twisting of the participant's entire trunk slowly and smoothly.

In the presence of intense resistance in any of these movements, the test should be stopped immediately.

Figure 1. Postures taken during the application of the Temporomandibular Dysfunction Assessment Test



All data evaluated were stored in excel spreadsheets, v2109, for later analysis. Statistical analyses were performed using SPSS software, version 23.0. For comparison purposes with the results of the TMD Assessment Test, this study recategorized the results of the FAIx, since the FAIx offers four possible answers and the TMD assessment test has a dichotomous outcome. The response no TMD in the FAIx remained the same in the recategorization, and all other possibilities (mild, moderate, and severe TMD) were recategorized as "Presence of TMD."

The chi-square test was performed to analyze the sensitivity (% of individuals with TMD) and specificity (% of individuals without the presence of TMD) in relation to the Fonseca Anamnesis Index, requesting the percentage of a combination of the evaluation test in relation to the FAIx. We also analyzed the positive predictive values (PPV) and negative predictive values (NPV), which means a prognostic value of, if the TMD test is positive or negative, the probability of the person having or not having TMD.

The calculation was made using the chi-square, with the PPV as reference the percentage of positive tests for the assessment test in relation to all positive tests, and for the NPV, the percentage of negative tests for the assessment test in relation to all negative tests. A statistical significance of 5% was considered for all tests applied.

Reproducibility was analyzed by means of the agreement between the three examiners simultaneously, using Fleiss' Kappa and Cohen's Kappa Index for 2x2 analysis. The Kappa index determines the proportion of observed and expected agreement, considering Kappa <0 (Poor), 0-0.2 (Weak), 0.21-0.4 (Fair), 0.41-0.6 (Moderate), 0.61-0.8 (Substantial), >0.81 (Near perfect).¹⁴

Results

The sample initially consisted of 13 individuals, 3 of whom were excluded due to trauma to the face. The final sample was composed of 10 individuals. The participants had a mean age of 26 ± 2.6 years, and five were female (50%), four were white (40%), eight were single (80%), five were physical therapy students (50%), seven participants (70%) consumed alcohol and practiced physical activity in each item, as shown in Table 1.

Table 1. Sociodemographic characteristics of the studied sample

Variables	n (%)
Gender	
Female	5 (50%)
Male	5 (50%)
Race	
White	4 (40%)
Brown	3 (30%)
Black	3 (30%)
Marital Status	
Single	8 (80%)
Married	1 (10%)
Divorced	1 (10%)
Occupation	
Student	5 (50%)
Physical therapist	4 (40%)
Psychologist	1 (10%)
Alcohol consumption	
Yes	7 (70%)
Physical Exercise	
Yes	7 (70%)

According to the Fonseca Anamnestic Index, eight individuals were diagnosed with the presence of TMD, while the TMD Assessment Test identified nine participants with the presence of this dysfunction.

With regard to diagnostic accuracy, the percentage of individuals diagnosed with the presence of TMD, which characterizes the sensitivity of the TMD Test was 100%, while the specificity value (or percentage of individuals diagnosed without TMD) was 50%. The Positive Predictive Value was 88% and the Negative Predictive Value was 50%, as shown in Table 2.

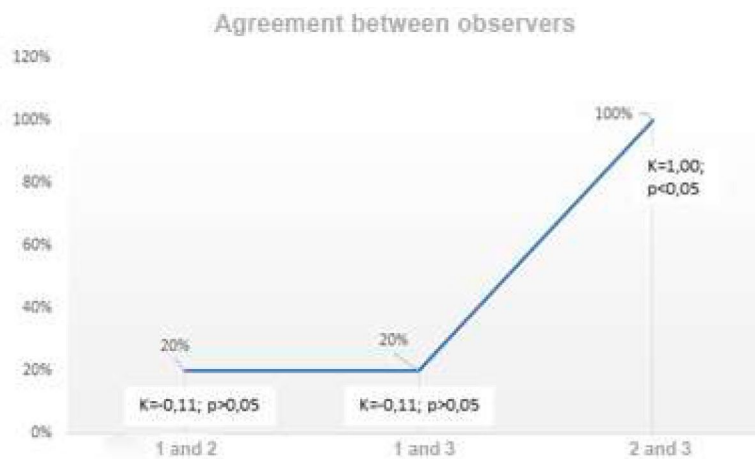
Table 2. Sensibility, Specificity, PPV and NPV of the TMD Test

	FONSECA ANAMNESIC INDEX		TOTAL	
	Presence of TMD	No presence of TMD		
TMD Assessment Test	Presence of TMD	8	1	9
	No presence of TMD	0	1	1
	TOTAL	8	2	10

Sensibility = $8/8 \times 100 = 100\%$; Specificity = $1/2 \times 100 = 50\%$; Positive Predictive Value (PPV) = $8/9 \times 100 = 88\%$; Negative Predictive Value (NPV) = $1/1 \times 100 = 100\%$

The analysis of overall agreement between the examiners Fleiss' Kappa test showed that there is reasonable reproducibility between results ($K = 0.26$ [95% CI: $-0.099 - 0.617$]; $z=1.420$; $p>0.05$), but not significant. Cohen's Kappa test showed poor reproducibility between observers 1 and 2 ($K = 0.11$, $p>0.05$) and between observers 1 and 3 ($K = 0.11$, $p>0.05$), and perfect reproducibility between observers 2 and 3 ($K = 1.00$, $p<0.05$).

Figure 2. Inter-rater accuracy



Discussion

The validation of diagnostic instruments is necessary to generate greater reliability and reproducibility of these tools, which corroborates the objective of validation and reproducibility of this study. Diagnostic accuracy studies aim to validate tests that accurately identify affected and unaffected individuals for a given dysfunction.

Compared to the FAIx, the TMD assessment test had perfect sensitivity, which means that the TMDAsT can diagnose TMD in people with TMD, a great tool used for diagnostic screening. However, the specificity of the TMDAsT was low, which means that the evaluative test can diagnose TMD in people who may not actually have the dysfunction, which is said to be a false positive test.

It is important to note that the difference in specificity between the tests was one participant. While the FAIx detected two participants without TMD, the TMDAsT detected only one, generating this result of 50%. That is, this proportion is overestimated and cannot function as a true measure of the test's specificity, requiring a larger sample.

Similar to the present study, a study conducted in 2001 to analyze the use of the Qualitative and Quantitative Orofacial Pain and TMD Screening Questionnaire found a sensitivity of 85.37% and specificity of 80% for patients with muscle disorders of the orofacial region (Kappa=0.45) and low sensitivity and specificity for intra-articular disorders (Kappa=0.043).⁵ Despite using an instrument that corresponds to the correct diagnosis of temporomandibular dysfunction, these authors indicate the need for a multidisciplinary evaluation for patients presenting specific signs and symptoms such as headaches, facial pain, and pain in the auricular pre-auricular region, and joint clicking.⁵

One study used the receiver operating characteristic (ROC) curve to determine the diagnostic accuracy, sensitivity, and specificity of the FAIx. Using the total score of this instrument (0 to 100 points) as an evaluation parameter, they found high diagnostic accuracy in the FAIx for TMD diagnosis, presenting a sensitivity of 86.30% and a specificity of 91.19%.¹¹

In 2018, authors would suggest avoiding the use of analgesics to select patients with TMD or candidates for intervention, as suggested in the present study as exclusion criteria and include clinical tests on the neck and shoulder girdle, thus taking analysis and consideration of postural assessment into the context of temporomandibular dysfunctions diagnosis, in order to promote an improvement in the RDC/TMD validation project.¹⁵

The low result in the test reproducibility may be associated with each examiner's ability to perform

the test since there was a difference in agreement between them. For example, we can see that examiners 2 and 3 agree with each other in the findings, but when each one was compared to rater 1, a greater difference was found between the results, suggesting that rater 1 disagrees with the others. With this, we can infer two possibilities: examiner 1 may be getting it wrong more than examiners 2 and 3, or examiners 2 and 3, despite they agree, are getting it wrong more than examiner 1.

The different results found in the examiners' diagnoses can be explained by the fact that the evaluative test for temporomandibular dysfunction is a subjective and examiner-dependent evaluation. Each person will perform the test according to their abilities, and the result will be what he or she could observe, which may have caused the difficulty of the three evaluators in agreeing on the diagnosis.

Since the diagnosis of temporomandibular dysfunction by the test depends on the evaluator noticing or not the presence of resistance during the performance of the proposed passive movements, another reason that could influence this divergence between the results could be the level of sensitivity training of the evaluators to determine a possible resistance in the body of the participants.

In a study conducted in 2009, the authors assessed the reliability of a form for diagnosing the severity of temporomandibular dysfunction, finding Kappa results between 0.725 and 0.838 for the proposed questions.¹⁶ In 2014, researchers validated and reproducibly a screening questionnaire for temporomandibular dysfunction in adolescents, finding Kappa values between 0.529 and 0.884.¹⁷ However, the two studies used a questionnaire or form as evaluation and performed the intra-examiner validation, while this research had as instrument a subjective evaluation test based on movements and performed the inter-examiner reproducibility.

In the present investigation, it is necessary to have more training time to perform the test since, due to the practice of the examiners, there is a divergence between the diagnoses, with Kappa ranging from poor to perfect. On the other hand, a study conducted in 2021, in which the Helkimo Index was validated and reproduced for the diagnosis of temporomandibular dysfunction, obtained inter-examiner reproducibility classified as moderate to substantial.¹⁸ However, although this study showed a high Kappa, the Helkimo Index, in addition to not assessing posture, was validated to assess the severity of signs and symptoms and did not provide diagnostic classification.⁹

In a study conducted in 2004, emphasis is given to the importance of posture assessment in patients with TMD. The temporomandibular joint is maintained in the orthostatic position by a complex mechanism involving the head, neck, and shoulder girdle muscles. Therefore, any change in their conformation can generate postural changes both in these specific locations and in other muscle chains.¹⁹ Thus, in 2016, a methodology measuring body posture and its relationship with temporomandibular dysfunction was validated and reproduced, evaluating deviations in the frontal and sagittal planes, with excellent inter-examiner results with Kappa > 0.87.²⁰ However, despite evaluating the association between TMD and posture, it is not a study that can be easily applied in clinical practice since it was evaluated by means of digital measurement with a specific methodology created to verify posture.

The limitations found in this study may be related to the context of the COVID-19 pandemic, in which the follow-up of this project underwent an interruption leading to interference in the sample size of the study and the training time of the raters. Thus, the authors consider the sample size a contributing factor to high sensitivity and low specificity value, thus requiring a larger number of subjects.

Conclusion

The TMD Assessment Test (TMDAsT) showed very high sensitivity, identified all individuals with TMD, and can be an excellent screening tool for TMD. However, as for reproducibility, this study points to the need for intense and constant training of examiners to perform the TMD evaluation test.

Author contributions

Neri YBO participated in the conception, design, search, and statistical analysis of the research data, data collection, results from interpretation, and scientific article writing. Francisco ACS, De Jesus IRN participated in the research design and data collection. Santos IEN participated in the conception, search, and statistical analysis of the research data, data collection, interpretation of results, writing of the scientific article. Ferreira RS participated in the conception and training of the evaluators. Góes ALB participated in the conception, design, orientation, and writing of the scientific article.

Competing interests

No financial, legal, or political conflicts involving third parties (government, private companies, and foundations, etc.) have been declared for any aspect of the submitted work (including but not limited to grants and funding, advisory board participation, study design, manuscript preparation, statistical analysis, etc.).

References

1. Costa LM, Medeiros DL, Ries LG, Beretta A, Noronha MA. Assessment of cross-cultural adaptations and measurement properties of self-report outcome measures in Portuguese relevant to temporomandibular disorders: a systematic review. *Fisioter Pesqui.* 2014;21(2):107-12. <https://doi.org/10.1590/1809-2950/35421022014>
2. Nunes AM, Martinez EM, Lopes PRR, Bittencourt MAV, Canedo PMDM. Association between the posterior muscle chain and the temporomandibular dysfunction severity. *Rev. Ciênc. Méd. Biol [Internet].* 2016;14(3):394. Available from: <http://repositorio.ufba.br/ri/handle/ri/23270>

3. Ferreira CLP, Silva MAMR, Felício CM. Signs and symptoms of temporomandibular disorders in women and men. *CoDAS*. 2016;28(1):17–21. <https://doi.org/10.1590/2317-1782/20162014218>
4. Cantero LS, Morales, JDT, Espinosa IA, Rodríguez ET. Temporomandibular disorders found in patients with malocclusions. *Rev Cubana Estomatol* [Internet]. 2013;50(4):374–87. Available from: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0034-75072013000400005&lng=es
5. Manfredi APS, Silva AAD, Vendite LL. The sensibility appreciation of the questionnaire for selection of orofacial pain and temporomandibular disorders recommended by the American Academy of Orofacial Pain. *Rev Bras Otorrinolaringol*. 2001;67(6):763–8. <https://doi.org/10.1590/S0034-72992001000600003>
6. Basso D, Corrêa E, Silva AM. Effect of global postural reeducation on body alignment and on clinical status of individuals with temporomandibular disorder associated to postural deviations. *Fisioter Pesqui*. 2010;17(1):63–8. <https://doi.org/10.1590/S1809-29502010000100012>
7. Azato FK, Castillo DB, Coelho TMK, Taciro C, Pereira PZ, Zomerfeld V, et al. Influence of temporomandibular disorders management on pain and global posture. *Rev. dor*. 2013;14(4):280–3. <https://doi.org/10.1590/S1806-00132013000400009>
8. Horak FB. Postural orientation and equilibrium: what do we need to know about neural control of balance to prevent falls?. *Age Ageing*. 2006;35 (Suppl 2):ii7–ii11. <https://doi.org/10.1093/ageing/afl077>
9. Chaves TC, Oliveira AS, Grossi DB. Main instruments for assessing temporomandibular disorders, part I: indices and questionnaires; a contribution to clinicians and researchers. *Fisioter Pesqui*. 2008;15(1):92–100. <https://doi.org/10.1590/S1809-29502008000100015>
10. Cunha SC, Nogueira RVB, Duarte ÂP, Vasconcelos BCE, Almeida RAC. Analysis of helkimo and craniomandibular indexes for temporomandibular disorder diagnosis on rheumatoid arthritis patients. *Braz J Otorrinolaryngol*. 2007;73(1):19–26. [https://doi.org/10.1016/s1808-8694\(15\)31117-4](https://doi.org/10.1016/s1808-8694(15)31117-4)
11. Peña G, Díaz W, Flores G, Marinkovic K, Romo F, Schulz R. Agreement between Diagnostic Criteria RDC/TMD and its update DC/TMD, applied to the inflammatory pathology of the Temporomandibular Joint. *Rev. Clin. Periodoncia Implantol. Rehabil*. 2019;12(2):70–3. <http://dx.doi.org/10.4067/S0719-01072019000200070>
12. Pastore GP, Goulart DR, Pastore PR, Prati AJ, Moraers M. Comparison of instruments used to select and classify patients with temporomandibular disorder. *Acta Odontol. Latinoam*. 2018;31(1):16–22. Cited: PMID: [30056462](https://pubmed.ncbi.nlm.nih.gov/30056462/).
13. Schwarzenbeck KCSB. Acurácia e correlação de instrumentos de avaliação da disfunção temporomandibular [dissertation] [Internet]. Campinas: Universidade Estadual de Campinas; 2013. Available from: http://bdtd.ibict.br/vufind/Record/CAMP_d0053ab521d7fd7f764a09b9f1cb322f
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–74. Cited: PMID: [843571](https://pubmed.ncbi.nlm.nih.gov/843571/)
15. Steenks M, TürpJ, de Wijer A. Reliability and Validity of the Diagnostic Criteria for Temporomandibular Disorders Axis I in Clinical and Research Settings: A Critical Appraisal. *J Oral Facial Pain Headache*. 2018;32(1):7–18. <https://doi.org/10.11607/ofph.1704>
16. Campos JADB, Gonçalves DAG, Camparis CM, Speciali JG. Reliability of a questionnaire for diagnosing the severity of temporomandibular disorder. *Rev. bras. fisioter*. 2009;13(1):8–43. <https://doi.org/10.1590/S1413-35552009005000007>
17. Franco-Micheloni AL, Fernandes G, Gonçalves DAG, Camparis CM. Temporomandibular disorders among Brazilian adolescents: reliability and validity of a screening questionnaire. *J. Appl. Oral Sci*. 2014;22(4):314–22. <https://doi.org/10.1590/1678-775720130694>
18. Alonso-Royo R, Sánchez-Torrel CM, Ibáñez-Vera AJ, Zagalaz-Anula N, Castellote-Caballero Y, Obrero-Gaitán E, et al. Validity and Reliability of the Helkimo Clinical Dysfunction Index for the Diagnosis of Temporomandibular Disorders. *Diagnostics (Basel)*. 2021;11(3):472. <https://doi.org/10.3390/diagnostics11030472>
19. Amantéa DV, Novaes AP, Campolongo GD, Barros TP. The importance of the postural evaluation in patients with temporomandibular joint dysfunction. *Acta Ortop Bras*. 2004;12(3):155–9. <https://doi.org/10.1590/S1413-78522004000300004>
20. Fernández RF, Carter P, Muñoz S, Silva H, Venegas GHO, Cantin M, et al. Evaluation of validity and reliability of a methodology for measuring human postural attitude and its relation to temporomandibular joint disorders. *Singapore Med J*. 2016;57(4):204–8. <https://doi.org/10.11622/smedj.2015159>