

FURTHER VALIDITY AND RELIABILITY OF TURKISH VERSION OF THE MANDIBULAR FUNCTIONAL IMPAIRMENT QUESTIONNAIRE (MFIQ) IN PATIENTS WITH TEMPOROMANDIBULAR DYSFUNCTION

Hasan Erkan Kılınc¹, Banu Ünver², Halil İbrahim Çelik³, Turgay Peyami Hocaoglu⁴

¹ Hacettepe University, Physical Therapy and Rehabilitation Faculty, Ankara, Turkey

² Lokman Hekim University, Health Science Faculty, Physiotherapy and Rehabilitation Department, Ankara, Turkey

³ Lokman Hekim University, Health Science Faculty, Physiotherapy and Rehabilitation Department, Ankara, Turkey

⁴ Lokman Hekim University, Faculty of Dentistry, Oral, Dental and Maxillofacial Surgery Department, Ankara, Turkey

ORCID: H.E.K. 0000-0002-6629-1619; B.U. 0000-0001-9758-6607; H.I.C. 0000-0003-3849-6382; T.P.H. 0000-0002-0281-8928

Corresponding author: Hasan Erkan Kılınc, **E-mail:** erkankilinc86@hotmail.com

Received: 18.03.2022; **Accepted:** 28.06.2022; **Available Online Date:** 31.01.2023

©Copyright 2021 by Dokuz Eylül University, Institute of Health Sciences - Available online at <https://dergipark.org.tr/en/pub/jbachs>

Cite this article as: Kılınc HE, Ünver B, Çelik HI, Hocaoglu TP. Further Validity and Reliability of Turkish Version of the Mandibular Functional Impairment Questionnaire (MFIQ) in Patients with Temporomandibular Dysfunction. J Basic Clin Health Sci 2023; 7: 214-222.

ABSTRACT

Purpose: The aim of this study is to develop a new Turkish version of MFIQ (MFIQ-T-N) by adding the cut off value and improving the validation to the Turkish version of the Mandibular Functional Impairment Questionnaire (MFIQ-T) in patients with Temporomandibular dysfunction.

Material and Methods: The psychometric properties of the MFIQ were examined in 242 participants diagnosed with TMD and 81 healthy participants. The internal construct validity was assessed with both exploratory and confirmatory factor analysis. External construct validity of the MFIQ-Tr-N was established by The Fonseca Anamnestic index (FAI), The Oral Health Impact Profile-14 (OHIP-14), and The Patient Health Questionnaire-9 (PHQ-9). Predictive validity was examined using the receiver operating characteristic (ROC) to validate the cut-off value of the MFIQ-T-N for diagnosis of TMD. For the determination of reliability internal consistency, test-retest reliability and smallest real difference (SRD95) were investigated.

Results: Factor analysis indicated that the MFIQ-T-N had three factors as masticatory, non-masticatory 1, and 2. The MFIQ-T-N was moderately correlated with the FAI (Spearman's rho = 0.553, p < .001) and OHIP-14 (rho = 0.534, p < .001). The cut-off point for MFIQ-T-N was set at >5 with 84% sensitivity and 87% specificity. The ICC and Cronbach's α values were found as 0.913 and 0.916, respectively. The SRD95 for Test-retest reliability was 7.004.

Conclusion: The MFIQ-T-N had sufficient psychometric features and was appropriate to use for participants with TMD. In addition, this new version provides researchers with additional opportunities to evaluate mandibular functions in the TMD patient group.

Keywords: Jaw, pain, temporomandibular disorders

INTRODUCTION

Temporomandibular dysfunction (TMD) is a complex problem associated with temporomandibular joint, masticatory muscles, and other related structures. It

is generally classified as joint origin, masticatory muscles origin and mixed type. The main symptoms of TMD are pain in mandibular region, tenderness with palpation in masticatory muscles, clicking sound

from temporomandibular joint, and restricted or asymmetrical mouth opening (1). The prevalence of TMD in the general population has been reported as 3-15% (2). TMD may lead to a wide spectrum of functional disabilities. Difficulty in speaking, laughing, yawning, and chewing is common in individuals with TMD. Ultimately, TMD has physiological, social, and psychological effects resulting in impaired quality of life (3).

Specific mandibular functions, such as mouth opening range and chewing ability, are evaluated by various techniques (4, 5). There are some anamnesis-based instruments for diagnosing temporomandibular dysfunction. Helkimo's Clinical Dysfunction Index, Fonseca Anamnestic Index, and Conti Anamnestic Index are of those anamnestic scales (6-8). In addition, to evaluate the functional outcomes caused by TMD, The Jaw Functional Limitation Scale (JFLS), which has both 8 and 20-item versions (9), and Temporomandibular Disability Index developed by Steigerwald et al. (10) are widely used. The Mandibular Function Impairment Questionnaire (MFIQ) is among the instruments which measure the functional impact of TMD and has been commonly used in clinical practice (11-14). This questionnaire is a valid and reliable scale that evaluates pain, jaw movements, and psychological impacts in TMD. MFIQ was originally developed in the English language. Portuguese version was conducted by Campos et al. (15) in 2012. In 2019 Chinese version was generated by Xu et al. (16). Yıldız et al. (17) conducted the first Turkish version study of the MFIQ in 2021. Kaynak et al. (18) conducted the Turkish validity and reliability study of the Fonseca Anamnestic Index (FAI).

The previous Turkish version study revealed that MFIQ was a valid and reliable instrument in Turkish patients with TMD (17). However, that study had a methodological shortcoming by not evaluating convergent validity. The fact that no cut-off score was determined is one of them. Therefore, the purpose of this study was to develop the current Turkish version of the MFIQ and to further investigate its psychometric properties (cross-cultural, criterion and construct validity, cut-off value, internal consistency, test-retest reliability, and measurement error) in patients with TMD.

MATERIAL AND METHODS

The study was approved by the ethics committee of Lokman Hekim University (approval number:

2021067). Informed consent form was obtained from all participants.

The Turkish translation and cultural adaptation of the MFIQ-T-N were completed using standard procedures established by Beaton et al. (19): 1) translation into Turkish by two native Turkish-speakers; 2) generation of a consensus version; 3) re-translation into English by professional translation experts; 4) comparison of this English scale with the original scale; 5) the creation of a semantically acceptable version for Turkish; and 6) finally, the last English version was approved by the MFIQ developers. Subsequently, the psychometric properties of the Mandibular Functional Impairment Scale-Turkish-New-Version (MFIQ-T-N) was evaluated according to the COSMIN criteria (20).

Participants

Between June 2021 and October 2021, 242 (214 female, 28 male) patients diagnosed with TMD with a mean age of 32.51 ± 9.70 years, and 81 (58 female, 23 male) healthy participants with a mean age of 29.8 ± 9.51 years participated in the study. Temporomandibular dysfunction was diagnosed by an experienced maxillofacial surgeon according to research diagnostic criteria for TMD (RDC/TMD) (21).

Instruments

Mandibular Function Impairment Questionnaire: The original version of the MFIQ was developed in 1993 by Stegenga et al. (11). The MFIQ consists of 17 items and 2 dimensions (D1: functional capacity, D2: feeding). Difficulty in performing various mandibular tasks is assessed with 5-point Likert scale (0 = no difficulty; 4 = very difficult or impossible without help). Higher scores indicate greater mandibular dysfunction. The sum score of the MFIQ ranges from 0 to 68 points.

Fonseca Anamnestic Index: FAI is a practical self-reported test that evaluates pain frequency, emotional stress, jaw function limitation, and different behavioral patterns associated with TMD (3). The FAI consists of ten questions with a three-point scale (0 = no, 5 = sometimes and 10 = yes). The total score ranges from 0-100 points, with higher scores indicating more severe symptoms. The Turkish version of this scale was generated by Kaynak et. al (18).

Oral health Impact Profile 14 (OHIP-14): OHIP-14 is the abbreviated/shortened form of the Oral Health Impact Profile, which was originally developed in

Table 1 Content validity ratio for the 17 items of the MFIQ-T-N

Items	Essential	Useful but and not -essential	Not-neccesary	CVR*
1 Social activities (family, friends, etc.)	23	2		0.84
2 Speaking	22	3		0.8
3 Taking large a bite	25			1
4 Chewing hard foods	25			1
5 Chewing soft food	25			1
6 Work and/or daily activities	23	2		0.84
7 Drinking	21	2	2	0.68
8 Laughing	24	1		0.92
9 Chewing elastic resistant food (elastic candies etc.)	25			1
10 Yawning	25			1
11 Kissing	18	2	5	0.44
12 A hard cookie	25			1
13 Meat (e.g beef).	25			1
14 A raw carrot	25			1
15 Baguette bread/white bread	25			1
16 Peanuts or almonds	25			1
17 An apple	25			1

*Minimum significant value according to Laewshe (1975), 0.42.(29)

Australia and consists of 49 questions. OHIP 14 evaluates oral health and its effects on the individual. OHIP-14 is evaluated with a 5-point Likert scale. Higher total score indicates worse oral health status and its effects on patients (22). Turkish version of the instrument was used in our study (23).

The patient health questionnaire-9 (PHQ-9): PHQ-9 evaluates nine symptoms of depression. The questionnaire contains nine questions scored between 0 (not at all) and 3 (nearly every day). The total score ranges from 0 to 27, and higher scores indicate an increased severity of depression (24). The Turkish version of this scale was generated by Sari et. al (25).

Content Validity

Seven physiotherapists and 18 dentists experienced in temporomandibular dysfunction examined the MFIQ items. As a result of this review, those experts classified the items. Finally, the content validity ratio (CVR) was calculated (26).

Construct Validity

The internal construct validity of the MFIQ-T-N was assessed by confirmatory factor analysis (CFA) and exploratory factor analysis (EFA). The sufficiency of the sample size and suitability of the data for the

factor analysis were examined by Kaiser Meyer Olkin (KMO) Test and Bartlett’s test, respectively. The factors were retained based on eigenvalues of more than one. To evaluate the factors’ goodness of fit, the ratio of the Chi-square test of model fit to the degrees of freedom (χ^2/df) [values of five or less], the Comparative Fit Index (CFI: >0.90 acceptable and

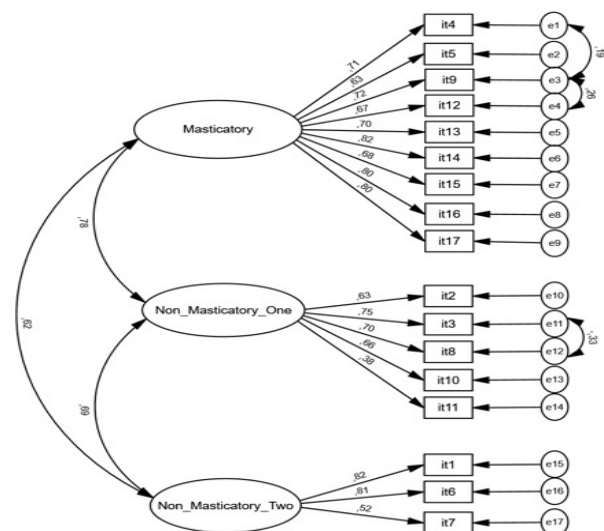


Figure 1. Three-factor model of the MFIQ-T-N

Table 2 Factor Loadings for the items of the MFIQ-T-N

Item	Factor loading		
	Factor 1	Factor 2	Factor 3
1. Social activities (family, friends, etc.)		,691	
2. Speaking			,447
3. Taking large a bite			,632
4. Chewing hard foods	,587		
5. Chewing soft food	,737		
6. Work and/or daily activities		,712	
7. Drinking		,850	
8. Laughing			,418
9. Chewing elastic resistant food (elastic candies etc.)	,721		
10. Yawning			,830
11. Kissing			,619
12. A hard cookie	,724		
13. Meat (e.g beef).	,763		
14. A raw carrot	,853		
15. Baguette bread/white bread	,779		
16. Peanuts or almonds	,693		
17. An apple (not cut into pieces)	,715		
Eigenvalues	7.467	1.499	1.096

Extraction method: principal component analysis. Rotation method: Oblimin with Kaiser Normalization.

>0.95 excellent), and the Root Mean Square Error of Approximation (RMSEA: <0.08 acceptable and <0.05 excellent) were used (27). External construct validity of the MFIQ-T-N was assessed by hypothesis testing (convergent validity). For the process of convergent validity, it was used the Spearman's correlation coefficients (r) for expected associations of the MFIQ-T with FAI, OHIP-14, and PHQ-9. validity, it was used the Spearman's correlation coefficients (r) for expected associations of the MFIQ-T with FAI, OHIP-14, and PHQ-9.

Criterion Validity

The evaluation of criterion validity of the MFIQ-T-N was examined through the process of predictive validity. ROC analyses were used to determine the cut-off score for the prediction of participants with TMD diagnosed by an experienced maxillofacial

surgeon according to RDC/TMD. The optimal cut-off score was identified using the Youden index (28).

Reliability

To determine the reliability of the MFIQ-T-N, internal consistency and test-retest reliability analyses were investigated. While the internal consistency was evaluated by the calculation of Cronbach's alpha coefficient >0.7. (29). Test-retest reliability was evaluated by intraclass correlation coefficient (ICC) with its confidence interval and spearman's correlation coefficients. The ICC values were interpreted in the following way: excellent agreement: >0.90; good agreement: 0.75–0.90; moderate agreement: 0.50–0.75. (30). The correlation coefficient values were interpreted in the following way: negligible: 0.00–0.30; low: 0.30–0.50; moderate:

0.50–0.70; high: 0.70–0.90 and very high: 0.90–1.00 (31).

To ensure the accuracy of the measurement method The Standard Error of Measurement (SEM) and the Smallest Real Difference (SRD) were calculated using the test-retest reliability statistics with the following formulae (32).

$$SEM = SD \times \sqrt{1 - ICC} \text{ and } SRD = SEM \times 1.96 \times \sqrt{2}$$

While IBM AMOS version 26.0 was used for confirmatory factor analyses, IBM SPSS Statistics for Windows version 26.0 was used for the remaining statistical analyses of this study. p value of <0.05 was considered statistically significant. A p value of <0.05 was accepted for statistical significance.

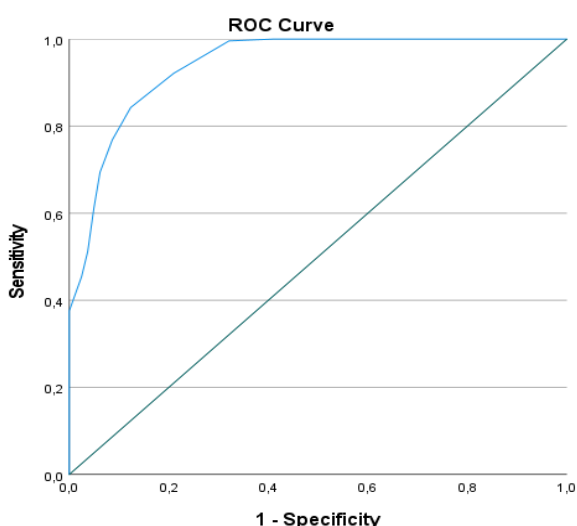


Figure 2. Illustration of the ROC curve

Ethical Consideration

The study was approved by the ethics committee of Lokman Hekim University (Date and approval number: 27.05.2021, 2021067).

RESULTS

Cross-cultural and content validity

The researchers and translators experienced some issues during the translation process. Since item 9 (chewing resistant food) is not fully understood in Turkey, it has been replaced with chewing elastic resistant food (elastic candies etc). Pork has a limited place in Turkish food culture, so item 13 (meat (e.g pork and beef)) option was changed to meat (e.g beef). Also, item 14 (French bread/White bread) has been changed to Baguette bread/White bread. Finally, the expert committee reached a consensus and verified the items as adequate by the following equivalences: semantic, idiomatic, experiential, and

conceptual. Additionally, as there was no individual with the lowest and highest scores, no ceiling and floor effects were observed for the. All items had a score above the CVR, therefore none of questionnaire items was not remove according to minimum significant value 0.42 (26). (Table 1).

Construct Validity

EFA was performed to determine the structure of the MFIQ-T-N. KMO and Bartlett’s sphericity test results revealed that the sample size was sufficient (KMO=0.914), and the items were appropriate (Bartlett’s test of sphericity: $\chi^2 = 2023.96, p < 0.001$) for the factor analysis. EFA indicated that the MFIQ-T-N had three factorial structures that accounted for 59.2 % of the total variance. The first factor consisted of 9 items (items 4, 5, 9, 12-17), which discerned masticatory function and accounted for 43.9% of the common variance. The second factor (8.8 % of the common variance) consisted of five items (items 2, 3, 8, 10, 11) which discern non-masticatory function 1, and the third factor (6.5 % of the common variance) consisted of three items (items 1, 6, 7) which discerned non-masticatory function 2 (Table 2).

The dimensional structure of the MFIQ-T-N obtained in the EFA was controlled by CFA. Modifications were made to optimize the dimensional structure of the scale according to the modification indices, which suggested adding a covariance between error items. The three-factor model (Figure 1) showed acceptable goodness-of-fit indices. ($X^2/df = 2.209, RMSEA = 0.071, CFI = 0.93$).

Hypothesis Testing- Convergent Validity

A correlation was found moderately between the MFIQ-T-N score and FAI ($p < 0.001, r = 0.553$) and OHIP 14 scores ($p < 0.001, r = 0.534$), while a weak correlation was found with PHQ-9 score ($p = 0.014, r = 0.158$) (Table 3).

Table 3 Correlation coefficients between the MFIQ-T-N and FAI, OHIP 14, PHQ-9

	FAI score	OHIP 14 score	PHQ-9 score
MFIQ-T-N total score	r=0.553	r=0.534	r=0.158
	p<0.001	p<0.001	p=0.014

Table 4 Results of the ROC analysis

	TMD group (n=242)	Control group (n=81)	p value	AUC	95% CI for AUC
MFIQ-T-N total score	X= 12.58 ±8.33	X= 2.38 ± 2.84	<0.001	0.941	0.911-0.971
	M=10 (7-16)	M=1 (0-4)			

AUC: Area under the curve, CI: Confidence interval, MFIQ-T: Mandibular Functional Limitation Scale-Turkish, X: Mean and standard deviation, M: Median and 25%-75% percentiles

Table 5 Reliability statistics for the total scale and the factors (subscales)

	Cronbach's α	Test-retest reliability ICC	Test-retest reliability rho (ρ)	SEM	SRD ₉₅
MFIQ-T-N total	0.916	0.913 (0.855-0.948)	0.780 (< 0.001)	2.527	7.004
Masticotary	0.909	0.913 (0.854-0.948)	0.805 (< 0.001)	1.672	4.634
Non-masticotary-1	0.743	0.858 (0.764-0.915)	0.765 (< 0.001)	1.002	2.777
Non-masticotary-2	0.755	0.615 (0.356-0.769)	0.572 (< 0.001)	0.819	2.270

ICC: Intraclass correlation coefficients (95% confidence interval), rho: Spearman's Correlations Coefficient, SEM: standard error of measurement, SRD₉₅: smallest real difference with 95% confidence

Criterion Validity- Predictive Validity

The clinical cut-off point for the MFIQ-T-N was set at >5 points with an area under the curve (AUC) of 0.941 (84% sensitivity and 87% specificity) (Table 4 and Figure 2)

Reliability

The Cronbach's alpha values for internal reliability were 0.916, 0.909, 0.743, and 0.755 for the MFIQ-T total, factor 1, 2 and 3, respectively. Based on these values, it was found that the total scale and the factor 1 and 2 had an excellent internal consistency, while the factor 2 and 3 had an acceptable internal consistency (Table 5). The ICC values with 95% confidence intervals for the MFIQ-T-N and factors (subscales) were given in Table 5. The test-retest reliability for the MFIQ-T-N total score and factor 1 was excellent while factor 2 and factor 3 were good and moderate, respectively. The test-retest reliability was also tested by Spearman correlation analysis. The high and moderate correlations between the scores indicated that MFIQ-T-N was reliable. In

addition, the SEM and SRD95 values for the total scale and the factors were also given in Table 5.

DISCUSSION

Current study was conducted to develop the Turkish version of the MFIQ and to investigate its psychometric properties in patients with TMD. The results indicated that new Turkish version of the MFIQ was valid and reliable for mandibular function in the Turkish population with TMD. The cross-cultural analysis enabled the modification of three items of the MFIQ-T-N, which was originally in English, so that the Turkish version could be easily understood. In the Chinese version, there was a modified item (French bread/white bread- donut) which was also modified in our translation. However, no items changes were required in the Portuguese version. In the previous Turkish version, only the French bread item was changed to bread (17). Differences in cultural adaptation may be due to the geographical diversity of the studies.

As a result of the content validity analysis of the Turkish version, there was no need to remove any

item. Likewise, no item was removed in previous Turkish version study (17). In the Chinese version, item 11 has been removed as it was below the minimum threshold. Although none of the items were below the CVR threshold value in the Portuguese version, four items were removed according to factor analysis.

When the construct validity of the new MFIQ Turkish version was examined, it was found that the three-factor structure was more suitable for Turkish population. Factor 1 included items on chewing, Factor 2 included social, work, and daily activities, Factor 3 included items on speaking, taking a large bite, laughing and kissing. MFIQ-T-N was designed as a three-factor structure in its original version, but since the number of items in the third factor is low and the consistency value is low, the second and third factors were combined (11). The previous Turkish, Chinese and Portuguese versions had a two-factor structure.

The clinical cut-off point for the MFIQ-T-N was set at > 5 points with AUC of 0.941 (84% sensitivity and specificity). Previous Turkish version study did not include a healthy group. Consequently, a ROC analysis was not performed, and a clinical cut-off score was not determined. In addition, the original, Chinese, and Portuguese versions of the MFIQ did not identify a clinical cut-off score. This study is unique in that it determines the cut-off value of the MFIQ.

The MFIQ-T-N total score was moderately correlated with FAI, OHIP-14, and weakly correlated with PHQ-9. FAI is an anamnesis-based questionnaire used to diagnose TMD, while MFIQ-T determines functional limitations of the mandible. On the other hand, OHIP-14 evaluates teeth, prosthesis, and general oral region health and its effects on psychological status and daily life. Our result shows that MFIQ-T reveals functional involvement in correlation with the TMD severity and oral health status. PHQ-9, examines the effects of general health issues on psychological and mental health. The weak correlation of this scale with the MFIQ-T suggests that mandibular dysfunction has a low effect on psychological and mental status. Previous version studies did not examine the correlation of the questionnaire with other relevant questionnaires.

The test-retest reliability was excellent both on a factorial and total basis in the previous Turkish, Portuguese, and Chinese version studies. Internal

reliability in MFIQ-T-N was excellent on a factor and total basis, similar to three studies.

CONCLUSION

Mandibular dysfunction seriously affects the quality of life. The MFIQ-T-N is an important tool in detecting existing functional problems in such patients. The new MFIQ-T-N has good reliability and validity for Turkish patients with TMD. Further evidence for convergent validity of the MFIQ in relation to a common clinical test (FAI and OHIP 19) in patients with TMD and important validity aspects, such as predictive validity (cut-off value) were added. Therefore, we believe that the self-reported mandibular movement limitation evaluation with MFIQ-T-N will be a clearer data source for directing the patient to both clinical and advance imaging evaluations. This situation is also important in terms of obtaining a multidirectional perspective for planning the rehabilitation of the patient.

Acknowledgement: The authors would like to thank all participants for their participation in the study.

Author contribution: Diagnosis: T.P.H., Concept: H.E.K., B.Ü., H.İ.Ç., Design: H.E.K., B.Ü., H.İ.Ç., Data Collection or Processing: H.E.K., B.Ü., Analysis or Interpretation: H.E.K., H.İ.Ç., Literature Search: H.E.K., T.P.H., Writing: H.E.K., B.Ü., H.İ.Ç., T.P.H.

Conflict of interests: No conflict of interest was declared by the authors.

Ethical approval: The study was approved by the ethics committee of Lokman Hekim University (Date and approval number: 27.05.2021, 2021067)

Funding The authors declared that this study received no financial support.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Laskin DM. Etiology of the pain-dysfunction syndrome. *The Journal of the American Dental Association.* 1969;79(1):147-53.
2. Freiwald HC, Schwarzbach NP, Wolowski A. Effects of competitive sports on temporomandibular dysfunction: a literature review. *Clin Oral Investig.* 2021;25(1):55-65.
3. Tjakkes GH, Reinders JJ, Tenverger EM, Stegenga B. TMD pain: the effect on health related quality of life and the influence of pain duration. *Health Qual Life Outcomes.* 2010;8:46(8):1-8.
4. Komagamine Y, Kanazawa M, Minakuchi S, Uchida T, Sasaki Y. Association between masticatory performance using a colour-changeable chewing gum and jaw movement. *J Oral Rehabil.* 2011;38(8):555-63.

5. Gallagher C, Gallagher V, Whelton H, Cronin M. The normal range of mouth opening in an Irish population. *J Oral Rehabil.* 2004;31(2):110-6.
6. Helkimo M. Studies on function and dysfunction of the masticatory system. *Acta Odontol Scand.* 1974;32(4):255-67.
7. Fonseca D, Bonfante G, Valle A, Freitas S. Diagnosis of the craniomandibular disfunction through anamnesis. *RGO.* 1994;42(1):23-8.
8. Conti PCA, Ferreira PM, Pegoraro LF, Conti JV, Salvador MC. A cross-sectional study of prevalence and etiology of signs and symptoms of temporomandibular disorders in high school and university students. *Journal of Orofacial pain.* 1996;10(3):253-62.
9. Ohrbach R, Larsson P, List T. The jaw functional limitation scale: development, reliability, and validity of 8-item and 20-item versions. *Journal of Orofacial pain.* 2008;22(3):41-47.
10. Steigerwald DP. The Steigerwald/Maher TMD disability questionnaire. *Today's Chiropractic.* 1997;26(4):86-91.
11. Stegenga B, de Bont LG, de Leeuw R, Boering G. Assessment of mandibular function impairment associated with temporomandibular joint osteoarthritis and internal derangement. *J Orofac Pain.* 1993;7(2):183-95.
12. Weinberg FM, Vermaire JA, Forouzanfar T, Rosenberg AJ, Speksnijder CM. Reproducibility and construct validity of the utrecht mixing ability Test to obtain masticatory performance outcome in patients with condylar mandibular fractures. *Journal of Oral Rehabilitation.* 2020;47(4):460-6.
13. Liu JN, Ong HS, Wang MY, Wan K, Qu XZ, Zhang CP. Randomized control trial comparing the titanium osteosynthesis and the biodegradable osteosynthesis in mandibulotomy access. *Head Neck.* 2019;41(4):915-23.
14. Pereira LJ, Steenks MH, de Wijer A, Speksnijder CM, van der Bilt A. Masticatory function in subacute TMD patients before and after treatment. *J Oral Rehabil.* 2009;36(6):391-402.
15. Campos JADB, Carrascosa AC, Maroco J. Validity and reliability of the Portuguese version of Mandibular Function Impairment Questionnaire. *Journal of Oral Rehabilitation.* 2012;39(5):377-83.
16. Xu L, Cai B, Fang Z. Translation and validation of a Chinese version of the Mandibular Function Impairment Questionnaire. *Journal of Oral Rehabilitation.* 2016;43(8):608-14.
17. Yildiz NT, Alkan A, Kulunkoglu BA. Validity and Reliability of the Turkish Version of Mandibular Function Impairment Questionnaire. *Cranio.* 2021;1(8):1-11.
18. Kaynak BA, Tas S, Salkin Y. The accuracy and reliability of the Turkish version of the Fonseca anamnestic index in temporomandibular disorders. *Cranio.* 2020;25(3):1-6.
19. Beaton D, Bombardier C, Guillemin F, Ferraz MB. Recommendations for the cross-cultural adaptation of health status measures. New York: American Academy of Orthopaedic Surgeons. 2002;12(2):1-9.
20. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research.* 2010;19(4):539-49.
21. Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders: review, criteria, examinations and specifications, critique. *J Craniomandib Disord.* 1992;6(4):301-55.
22. Allen F, Locker D. A modified short version of the oral health impact profile for assessing health-related quality of life in edentulous adults. *Int J Prosthodont.* 2002;15(5):446-50.
23. Basol ME, Karaagaclioğlu L, Yillmaz B. Türkçe Ağız Sağlığı Etki Ölçeğinin Geliştirilmesi-OHIP-14-TR. *Türkiye Klinikleri Dishekimliği Bilimleri Dergisi.* 2014;27(2):4620-2.
24. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606-13.
25. Sari YE, Kokoglu B, Balcioglu H, Bilge U, Colak E, Unluoglu I. Turkish reliability of the patient health questionnaire-9. *Biomedical Research-India.* 2016;27(2):460-2.
26. Lawshe CH. A quantitative approach to content validity. *Personnel Psychology.* 1975;28(4):563-75.
27. Geldhof GJ, Preacher KJ, Zyphur MJ. Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods.* 2014;19(1):72.
28. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. *Indian Pediatr.* 2011;48(4):277-87.

29. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297-334.
30. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155-63.
31. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J*. 2012;24(3):69-71.
32. Stratford PW. Estimating the standard error of measurement from reliability studies. *Physiother Can*. 2004;56(1):27-30.