Reliability of a questionnaire for diagnosing the severity of temporomandibular disorder

Confiabilidade de um formulário para diagnóstico da severidade da disfunção temporomandibular

Campos JADB[1], Gonçalves DAG[2], Camparis CM[3], Speciali JG[4]

Abstract

Background: Temporomandibular disorder (TMD) is a collective term that encompasses many clinical problems involving the masticatory muscles, temporomandibular joints (TMJ) and associated structures and it has high prevalence among populations. **Objectives:** Because Brazilian studies have used the instrument proposed by Da Fonseca et al. (1994) to diagnose the severity of TMD, this study was conducted to investigate and estimate the internal consistency and reproducibility of this method. **Methods:** We used a probability sampling design to select 1230 participants over the age of 18 years who were living in the city of Ribeirão Preto, SP, Brazil. The interviews were conducted by a single interviewer over the phone. The internal consistency was analyzed by calculating the Kuder-Richardson coefficient (kr-20), and kappa statistics (κ) were used to estimate the reproducibility. **Results:** The internal consistency of the questionnaire was 0.5594, thus indicating that validation was lower than desired. Questions 1, 2, 3, 6 and 7 had greater contribution towards the total kr-20 coefficient, and the consistency of the instrument was higher when it was composed only of these questions (0.7044). "Good" and "Excellent" reproducibility was observed for these same questions. **Conclusions:** Based on these data, it is suggested that the questionnaire proposed by Da Fonseca et al. (1994) should be adapted to include only questions 1, 2, 3, 6 and 7 of the initial version. This would help improve the reliability of the instrument. The need for validation studies must also be emphasized to ensure that the new version of the instrument has adequate psychometric characteristics.

Key words: orofacial pain; temporomandibular joint; diagnosis; reliability.

Da Fonseca DM, Bonfante G, Valle AL, Freitas SFT. Diagnóstico pela anamnese da disfunção craniomandibular. Revista Gaúcha de Odontologia. 1994;4(1):23-2.

Resumo

Contextualização: Disfunção temporomandibular (DTM) é um termo coletivo que engloba vários problemas clínicos envolvendo a musculatura da mastigação, as articulações temporomandibulares (ATM) e suas estruturas associadas, com alta prevalência nas populações. **Objetivos:** Sabendo-se que estudos brasileiros vêm utilizando o instrumento proposto por Da Fonseca et al. (1994) para diagnóstico da severidade desta disfunção, realizou-se este estudo com o objetivo de verificar e estimar a consistência interna e a reprodutibilidade do mesmo. **Métodos:** O delineamento amostral adotado foi o probabilístico, e participaram 1230 indivíduos moradores da cidade de Ribeirão Preto (SP), maiores de 18 anos de idade. As entrevistas foram realizadas por um único entrevistador por meio de ligações telefônicas. Para estudo da consistência interna, calculou-se o Coeficiente de Kuder-Richardson (kr-20) e para estimar a reprodutibilidade, utilizou-se a estatística Kappa (κ). **Resultados:** A consistência interna do formulário foi de 0,5594, apontando para uma validação abaixo do desejado. Observou-se maior contribuição das questões 1, 2, 3, 6 e 7 para o coeficiente kr-20 total e maior consistência do instrumento quando composto apenas pelas mesmas (0,7044). Observou-se reprodutibilidade "Boa" e "Ótima" para as questões. **Conclusões:** Frente ao exposto, sugere-se que o formulário proposto por Da Fonseca et al. (1994) seja adaptado, ficando composto apenas pelas questões 1, 2, 3, 6 e 7 da versão inicial, colaborando, assim, para aumento da confiabilidade do instrumento. Deve-se ressaltar ainda a necessidade da realização de estudos de validade para assegurar adequadas características psicométricas à nova versão do instrumento.

Palavras-chave: dor orofacial; articulação temporomandibular; diagnóstico; confiabilidade.

Da Fonseca DM, Bonfante G, Valle AL, Freitas SFT. Diagnóstico pela anamnese da disfunção craniomandibular. Revista Gaúcha de Odontologia. 1994;4(1):23-2.

[1] Departament of Social Dentistry, Faculdade de Odontologia de Araraquara (FOAR), Universidade Estadual Paulista (UNESP), Araraquara (SP), Brazil

[2] Graduate Program in Oral Rehabilitation, FOAR, UNESP

[3] Departrament of Dentistry Materials and Prostheses, FOAR, UNESP

[4] Departament of Neurology, Psychiatry and Medical Psychology, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo (USP), Ribeirão Preto (SP), Brasil

Correspondence to: Juliana Alvares Duarte Bonini Campos, Faculdade de Odontologia de Araraquara, Departamento de Odontologia Social, Rua Humaitá, 1.680, Centro, CEP14801-903, Araraquara (SP), Brazil, e-mail: jucampos@foar.unesp.br

Introduction ::::.

Temporomandibular disorder (TMD) is a collective term which embraces several clinical problems involving the masticatory muscles, temporomandibular joints (TMJ) and its associated structures or both. These changes are characterized mainly by pain in the temporomandibular area or in the masticatory muscles, limitations or deviations in the jaw movement, and TMJ sounds during the jaw function[1].

The international literature includes a great amount of tracer studies on TMD. The results of these studies vary considerably due to the different methodologies applied, to the characteristics of the populations and to the data collection[1-7].

According to several studies published in different countries, TMD has shown a high prevalence in the populations[2,8-11]. Agerberg and Inkapoo[8] verified that 88.0% of the 637 individuals showed signs and symptoms of TMD. Carlsson[4] observed a TMD prevalence varying between 6.0% and 93.0%. Pedroni, Oliveira and Guaratini[9] applied the Anamnesis Index (AI) proposed by Da Fonseca et al.[2] and observed that 68.0% of the sample of college students demonstrated some degree of TMD. Gesch et al.[10] observed a TMD prevalence of 51.8% in a sample composed by 4,289 adult individuals. In a sample composed by Brazilian students, Bevilaqua-Grossi et al.[11] observed that 87.0% of the individuals demonstrated some degree of TMD.

Several instruments for TMD diagnosis have been presented in the literature, but there is no consensus diagnostic criterion. Dworkin and Leresche[12] proposed the Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) due to the need to use an instrument which is universally accepted and validated. This instrument identifies the complex interaction between the physical and psychological dimensions of chronic pain and it develops into a double axis system in an attempt to provide a reliable measurement of TMD signs and symptoms (Axis I), as well as theist associated psychological and psychosocial factors (Axis II).

However, this is quite an extensive instrument which requires the presence of the individual for the TMD diagnosis. Therefore, the use of the RDC/TMD in great epidemiological studies can become unviable depending on the interview technique and/or on the available time for data collection.

The AI proposed by Da Fonseca et al.[2] is one of the available instruments in Portuguese language for the characterization of TMD symptoms and it was developed to classify patients according to the severity of those symptoms. These authors studied its correlation with the Helkimo[13,14] modified Clinical Index (CI) and obtained a high positive correlation (r=0.95). Based on these results, they recommended that public health and screening services should use the AI to trace TMD sufferers due to its simplicity, speed and low cost even though it is not a diagnostic instrument of TMD as the RDC/TMD. Thus, the AI would act as a preliminary TMD tracing tool, and after the affected population was identified, a more thorough investigation would be conducted with the use of diagnostic instruments including a physical exam. However, although several studies have used the AI[1,9], the reliability of this instrument to detect the presence or absence of TMD has not been investigated.

Thus, the present study was conducted with the purpose of estimating the internal consistence and reproducibility of the questionnaire proposed by Da Fonseca et al.[2] for the diagnosis of TMD severity.

Methods ::::.

Participants

The sample design adopted was the stratified random sampling. The strata were based on the census area and sex. The sample size was established through the sampling process for finite population. Based on the data showed on the American Academy of Orofacial Pain (AAOP), the value admitted for the expected TMD prevalence in the population was set at 40%. The α level considered for the analyses was set at 0.05 and the sampling error was set at 3%. Considering a 20% rate of absenteeism, the minimum sample size was estimated at 1,230 individuals to achieve a power of 90%. Residents of the city of Ribeirão Preto (SP), over the age of 18, were randomly selected from the telephone directory to take part in this study.

Instrumentation

The questionnaire proposed by Da Fonseca et al.[2] consists of 10 questions that allow the answers "Yes", "Sometimes" and "No", with scores of 10, 5 and zero, respectively. The sum of the points classifies the interviewees in the following TMD categories: Absent (0 to 15 points), Mild (20 to 40 points), Moderate (45 to 60 points) and Severe (70 to 100 points).

Based on the assertion[15-18] that TMD signs and symptoms can be transitory and self-limiting over time and generally characterized as intermittent, "yes" and "sometimes" were considered positive answers to the existence of a sign or symptom, as previously suggested by Bevilaqua-Grossi et al.[11]. Therefore, in the present study we opted to use a dichotomous scale (yes or no).

Procedures

The option of conducting the interviews over the phone was based on the literature[19-23] taking into account that the selected instrument is simple. The duration of interviews was around ten minutes, which is less than what is suggested by Rea and Parker[24] for a phone interview.

When the individual answered the phone, the researcher identified himself and read a statement informing the resident about the objective and content of the study and pointed out the confidentiality of the details. Thus, the individual was able to choose whether or not to take part in the study. This procedure for participation consent was based on the advice for data collection by telephone from the Research Ethics Committee of Hospital das Clínicas de Faculdade de Medicina de Ribeirão Preto of Universidade de São Paulo (protocol nº 4136/2002). When a call was not answered or the individual opted not to participate, a new phone number was randomly selected. Thus, there was no sample loss. The phone calls were made by a single trained examiner (κ=0.92).

Statistical analyses

The internal consistency of the instrument was evaluated using the Kuder-Richardson Coefficient (kr-20). The reliability of each question was verified through the kr-20 calculation when the referred question was excluded. If the obtained coefficient was lower than the test's kr-20 coefficient, this would indicate that the question was good as it increased the test reliability and vice-versa.

The calculated kr-20 score was compared with the threshold usually considered adequate (kr≥0.70)[25]. The mean correlation between items was also observed. According to Streiner and Norman[26], this correlation must be higher than 0.20 for questions to be considered adequate components of the instrument.

Forty individuals were selected in a draw to answer the questionnaire for the second time after 7 days. Kappa statistics (κ)[27] were used to verify the reproducibility of the instrument's questions. Landis and Koch[28] standards were used to classify the agreement scores. These patterns are shown in Table 1.

Table 1. Kappa value (κ) and its classification according to Landis and Koch[28].

κ	Agreement
<0.00	Poor
0.00 ⊢ 0.21	Slight
0.21 ⊢ 0.41	Fair
0.41 ⊢ 0.61	Moderate
0.61 ⊢ 0.81	Substantial
0.81 ⊢ 1.00	Almost perfect

The present study was conducted after the approval of the Research Ethics Committee of Hospital das Clínicas of Faculdade de Medicina de Ribeirão Preto (FMRP/USP) (protocol nº 4136/2002).

Results

The sample was composed by 599 female and 631 male individuals varying in age from 15 to 64 years, with a mean age of 38.3±13.76 years. Among the participants, 278 (22.60%) had no TMD; 770 (62.60%) had mild TMD; 126 (10.24%) had moderate TMD and 56 (4.56%) had severe TMD.

The internal consistency of the questionnaire proposed by Da Fonseca et al.[2], evaluated by the kr-20 coefficient, was 0.5594. This consistency was lower than desired. The reliability of each question is shown in Table 2.

Although all of the questions demonstrated a reliability level below what was desired, questions 1, 2, 3, 6 and 7 stand out in Table 2 for the total kr-20 coefficient. When these questions were eliminated, the coefficient had a significant decrease, which indicates their great contribution to the reliability of the instrument. Thus, for the new internal consistency analysis of the instrument, the other questions were discarded and the questionnaire consisted of only five questions (Table 3).

After the adjustment, there was an important increase in the questionnaire's internal consistency (0.7044). With this kr-20 coefficient score, the questionnaire could now be considered a consistent method. Another aspect to be pointed out is the adequate mean correlation between questions.

The intra-examiner reproducibility, verified by the κ statistics of each question, is demonstrated in the Table 4. The questions demonstrated "Good" and "Excellent" reproducibility, which confers an adequate agreement level to the method.

Discussion

The questionnaire proposed by Da Fonseca et al.[2] has been used in epidemiological studies of TMD. Therefore the accuracy of this measurement method should be studied to allow the verification of the instrument's capacity to evaluate the intended object. In this sense, the reliability analysis is an extremely important stage and it can be accomplished through the calculation of the internal consistency and reproducibility of the instrument.

The questionnaire initially demonstrated a kr-20 of 0.5094 (Table 2), i.e. it was only able to measure 50.94% of the real impact of TMDs on the studied population. When questions 4, 5, 8, 9 and 10 were excluded, this value increased to 70.44%

Table 2. Kuder-Richardson (kr-20) coefficient for each item of the instrument proposed by Da Fonseca et al.[2]. Ribeirão Preto, 2004.

Questions	Mean inter-item correlation	kr-20
1. Do you have difficulty opening your mouth wide?	0.0883	0.4659
2. Do you have difficulty moving your jaw to the sides?	0.0918	0.4763
3. Do you feel fatigue or muscle pain when you chew?	0.0985	0.4957
4. Do you have frequent headaches?	0.1184	0.5471
5. Do you have neck pain or a stiff neck?	0.1203	0.5518
6. Do you have ear aches or pain in that area (TMJ)?	0.1037	0.5101
7. Have you ever noticed any noise in your TMJ while chewing or opening your mouth?	0.1074	0.5198
8. Do you have any habits such as clenching or grinding your teeth?	0.1227	0.5573
9. Do you feel that your teeth do not come together well?	0.1395	0.5934
10. Do you consider yourself a tense (nervous) person?	0.1361	0.5864
Scale Test	0.1127	0.5594

Table 3. Kuder-Richardson (kr-20) coefficient for each item of the adapted instrument by Da Fonseca et al.[2]. Ribeirão Preto, 2004.

Questions	Mean inter-item correlation	kr-20
1. Do you have difficulty opening your mouth wide?	0.2856	0.6153
2. Do you have difficulty moving your jaw to the sides?	0.2979	0.6292
3. Do you feel fatigue or muscle pain when you chew?	0.3212	0.6544
6. Do you have ear aches or pain in that area (TMJ)?	0.3506	0.6835
7. Have you ever noticed any noises in your TMJ while chewing or opening your mouth?	0.3585	0.6909
Scale Test	0.3228	0.7044

Table 4. Kappa statistics (κ) and its classification for each item of the adapted instrument by Da Fonseca et al.[2]. Ribeirão Preto, 2004.

Questions	κ	Agreement
1. Do you have difficulty opening your mouth wide?	0.787	Substantial
2. Do you have difficulty moving your jaw to the sides?	0.725	Substantial
3. Do you feel fatigue or muscle pain when you chew?	0.771	Substantial
6. Do you have ear aches or pain in that area (TMJ)?	0.805	Substantial
7. Have you ever noticed any noises in your TMJ while chewing or opening your mouth?	0.838	Almost perfect

(Table 3), which shows a considerable improvement in the identification of TMDs. This score change can be attributed to the fact that these questions were not consistent enough to compose the instrument. Thus, we suggest a questionnaire composed only by questions 1, 2, 3, 6, 7.

These five questions are among the ones suggested by the AAOP for the initial TMD screening. In agreement with AAOP recommendations, a positive answer to at least one of these questions can justify a more comprehensive evaluation if it represents a concern to the patient or if it is considered as clinically significant by the physician[1].

Lobezzoo and Lavigne[29] also point out limitations in the AI and suggest that questions 8 and 10, which intend to evaluate parafunctional habits and stress perception, respectively, are not good predictors of the diagnosis of TMD severity. In contrast, Bevilaqua-Grossi et al.[11] considered that questions 4 and 5, which intend to evaluate the headache and nape pain or torticollis, respectively, are also poor predictors of the TMD severity.

After the reorganization of the questionnaire, the reproducibility study was conducted. The literature suggests, for a reproducibility study, an interval of up to one month between the two applications; however, it should be pointed out that the reproducibility should be evaluated within the context of the study.

The TMD investigation requires a significant decrease in the interval between the measures due to the intermittent characteristic of its signs and symptoms, leading to periods of improvement and even remission of the disorder. For that reason, an interval of just one week was chosen for this study. In Table 4, we see that all the questions demonstrated an appropriate reproducibility.

Considering that epidemiological and clinical TMD studies are subject to several errors associated mainly with methodological aspects[7,30], it is very important to evaluate the internal consistency and the reproducibility of instruments in order to make a correct TMD diagnosis.

Therefore, it is suggested that the questionnaire proposed by Da Fonseca et al.[2] should not be discarded as it possesses useful characteristics for epidemiological studies such as simplicity, speed, low cost and the possibility of data collection over the telephone. However, the challenge of obtaining a reliable and valid instrument for TMD tracing in populations must still be met, and the suggestion of the AI adaptation put forward by this study is only the first step in the search for an appropriate instrument as it is still necessary to study the validity of this adaptation.

Conclusion :::.

Based on the exposed results, it is suggested that the questionnaire proposed by Da Fonseca et al.[2] should be adapted to include only questions 1, 2, 3, 6 and 7 of the initial version which will collaborate to the increase of the instrument's reliability. However, the need for validation studies must be emphasized to ensure that the new version of the instrument has adequate psychometric characteristics.

References :::.

1. Okeson JP, editor. Dor Orofacial – Guia para Avaliação, Diagnóstico e Tratamento. Academia Americana de Dor Orofacial. São Paulo: Quintessence Editora Ltda; 1998.

2. Da Fonseca DM, Bonfante G, Valle AL, de Freitas SFT. Diagnóstico pela anamnese da disfunção craniomandibular. Rev Gauch de Odontol. 1994;4(1):23-32.

3. Dworkin SF, Huggins KH, LeResche L, Vankorff M, Howard J, Truelove E, et al. Epidemiology of signs and symptoms in temporomandibular disorders: clinical signs in cases and controls. J Am Dent Assoc. 1990;120(3):273-81.

4. Carlsson GE. Epidemiology and treatment need for temporomandibular disorders. J Orofac Pain. 1999;13(4):232-7.

5. de Lucena LBS, Kosminsky M, da Costa LJ, Góes PS. Validation of the Portuguese version of the RDC/TMD Axis II questionnaire. Braz Oral Res. 2006;20(4):312-7.

6. Nilsson IM, List T, Drangsholt M. The reliability and validity of self-reported temporomandibular disorder pain in adolescents. J Orofac Pain. 2006;20(2):138-44.

7. Campos JADB, Carrascosa AC, Loffredo LCM, Faria JB. Consistência interna e reprodutibilidade da versão em português do critério de diagnóstico na pesquisa para desordens temporomandibulares (RDC/TMD – Eixo II). Rev Bras Fisioter. 2007;11(6):451-9.

8. Agerberg G, Inkapöol I. Craniomandibular disorders in urban Swedish population. J Craniomandib Disord. 1990;4(3):154-64.

9. Pedroni CR, de Oliveira AS, Guaratini MI. Prevalence study of signs and symptoms of temporomandibular disorders in university students. J Oral Rehabil. 2003;30(3):283-9.

10. Gesch D, Bernhardt O, Alte D, Schwahn C, Kocher T, John U, et al. Prevalence of signs and symptoms of temporomandibular disorders in a urban and rural German population: Results of a population-based Study of Health in Pomerania. Quintessence Int. 2004;35(2):143-50.

11. Bevilaqua-Grossi D, Chaves TC, Oliveira AS, Monteiro-Pedro V. Anamnestic Index severity and signs and symptoms of TMD. J Cranio Practice. 2006;24(2):112-8.

12. Dworkin SF, LeResche L. Research diagnostic criteria for temporomandibular disorders: review, criteria, examinations and specifications, critique. J Carniomandib Disord. 1992;6(4):301-55.

13. Helkimo M. Studies on function and dysfunction of the masticatory system. II – Index for anamnestic and clinical dysfunction and oclusal state. Sven Tadlak Tidskr. 1974;67(2):101-21.

14. Helkimo M. Studies on function and dysfunction of the masticatory system. III – Analyses of anamnestic and clinical recordings of dysfunction with the aid of indices. Sven Tadlak Tidskr. 1974;67(3):165-81.

15. Greene CS, Laskin DM. Long term evaluation of treatment for myofascial pain-dysfunction syndrome: a comparative analysis. J Am Dent Assoc. 1983;107(2):235-8.

16. Mejersjö C, Carlsson GE. Long-term results of treatment for temporomandibular pain-dysfunction. J Prosthet Dent. 1983;49(6):809-15.

17. Fricton JR. Recent advances in temporomandibular disorders and orofacial pain. J Am Dent Assoc. 1991;122(10):24-32.

18. Magnusson T, Carlsson GE, Egermark I. Changes in subjective symptoms of craniomandibular disorders in children and adolescents during a 10-year period. J Orofac Pain. 1993;7(1):76-82.

19. Duckro PN, Tait RC, Margolis RB, Deshields TL. Prevalence of temporomandibular symptoms in a large United States metropolitan area. Cranio. 1990;8(2):131-8.

20. Glass EG, McGlynn FD, Glaros AG, Melton K, Romans K. Prevalence of temporomandibular disorders symptoms in a major metropolitan area. Cranio. 1993;11(3):217-20.

21. Goulet JP, Lavigne GJ, Lund JP. Jaw pain prevalence among French-speaking Canadians in Quebec and related symptoms of temporomandibular disorders. J Dent Res. 1995;74(11):1738-44.

22. Locker D, Slade G. Prevalence of symptoms associated with temporomandibular disorders in a Canadian population. Community Dent Oral Epidemiol. 1988;16(5):310-6.

23. Galan I, Rodríguez-Artalejo F, Zorilla B. Reproducibilidad de um cuestionario telefónico sobre factores de riesgo asociados al comportamiento y las prácticas preventivas. Gac Sanit. 2004;18(2):118-28.

24. Rea LM, Parker RA. Metodologia de Pesquisa – do planejamento à execução. São Paulo: Pioneira Thomson Learning Ltda; 2000.

25. Nunnally JC, Bernstein IH. Psychometric Theory. 3rd ed. New York: WCB/McGraw-Hill; 1994.

26. Streiner D, Norman G. Health measurement scales. A practical guide to their development and use. Oxford: Oxford University Press; 1995.

27. Light RJ. Measures of response agreement for qualitative data: some generalizations and alternatives. Psychol Bull. 1971;76:365-77.

28. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.

29. Lobbezoo F, Lavigne GJ. Do bruxism and temporomandibular disorders have a cause-and-effect relationship? J Orofac Pain. 1997;11(1): 15-23.

30. Fricton JR, Schiffman EL. Reliability of a craniomandibular index. J Dent Res. 1986;65(11):1359-64.