www.scielo.br/jaos
http://dx.doi.org/10.1590/1678-775720130694

Temporomandibular disorders among Brazilian adolescents: reliability and validity of a screening questionnaire

Ana Lucia FRANCO-MICHELONI[1,2], Giovana FERNANDES[1], Daniela Aparecida de Godoi GONÇALVES[1], Cinara Maria CAMPARIS[1]

1- Department of Dental Materials and Prosthodontics, Araraquara Dental School, Univ. Estadual Paulista - UNESP, Araraquara, SP, Brazil.
2- Centro Universitário de Araraquara - UNIARA, Araraquara, SP, Brazil.

Corresponding address: Ana Lúcia Franco-Micheloni - Departamento de Materiais Odontológicos e Prótese - Faculdade de Odontologia de Araraquara - Rua Humaitá, 1680 - Centro - 14801-903 - Araraquara - SP - Brazil - Phone: 55-16-33016412 Fax: 55-16-33016406 - e-mail: analu.franco@hotmail.com

Submitted: December 9, 2013 - Modification: February 23, 2014 - Accepted: March 31, 2014

ABSTRACT

Temporomandibular disorders (TMD) screeners assume significant item overlap with the screening questionnaire proposed by the American Academy of Orofacial Pain (AAOP). Objective: To test the reliability and validity of the Portuguese version of AAOP questions for TMD screening among adolescents. Material and Methods: Diagnoses from Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) Axis I were used as reference standard. Reliability was evaluated by internal consistency (KR-20) and inter-item correlation. Validity was tested by sensitivity, specificity, predictive values, accuracy and receiver operating characteristic (ROC) curves, the relationship between the true-positive rate (sensitivity) and the false-positive rate (specificity). Test-retest reliability of AAOP questions and intra-examiner reproducibility of RDC/TMD Axis I were tested with kappa statistics. Results: The sample consisted of 1307 Brazilian adolescents (56.8% girls; n=742), with mean age of 12.72 years (12.69 F/12.75 M). According to RDC/TMD, 397 [30.4% (32.7% F/27.3% M)] of adolescents presented TMD, of which 330 [25.2% (27.6% F/22.2% M)] were painful TMD. Because of low consistency, items #8 and #10 of the AAOP questionnaire were excluded. Remaining items (of the long questionnaire version) showed good consistency and validity for three positive responses or more. After logistic regression, items #4, #6, #7 and #9 also showed satisfactory consistency and validity for two or more positive responses (short questionnaire version). Both versions demonstrated excellent specificity (about 90%), but higher sensitivity for detecting painful TMD (78.2%). Better reproducibility was obtained for the short version (k=0.840). Conclusions: The Portuguese version of AAOP questions showed both good reliability and validity for the screening of TMD among adolescents, especially painful TMD, according to RDC/TMD.

Keywords: Temporomandibular joint disorders. Questionnaires. Validation studies. Adolescent.

INTRODUCTION

Advances in the understanding regarding the prevalence, etiology and natural progression for temporomandibular disorders (TMD), as well as for the establishment of their treatment strategies, are dependent on reliable and valid diagnostic criteria[26].

Indeed, there has been notable development of screeners for both clinical and epidemiological purposes. The American Academy of Orofacial Pain (AAOP) published its parameters in 1990. Since the first edition of the screening questionnaire was presented, subsequent editions have been republished over the course of years, but no significant changes have been made until the fifth edition[2] published in 2013. Although the AAOP questions resulted from consensus among experts, and evolved from instruments and protocols presented in other publications, they have never been sufficiently tested in adults or adolescents.

These questions were proposed for the initial screening of patients, and the number of positive responses would help clinicians decide whether a more comprehensive evaluation would be necessary to obtain a definitive TMD diagnosis.

The Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD)[6] were presented in 1992 to promote standardization and replication of research on TMD. They enable clinical researchers to classify TMD subtypes in a similar manner in different countries and languages. The RDC/TMD have been tested in many studies and, although their qualities of accuracy, validity and reliability have been demonstrated, a new version has recently been proposed[25]. They have been applied in countless researches; however, they are difficult to apply in large epidemiological studies, because the protocols are long, time-consuming and require a face-to-face evaluation[3,21].

A few TMD screeners have been proposed to assess TMD symptoms in epidemiologic studies with children, adolescents[21,29] and adults[12,30]. All these screeners assume significant item overlap with the ten questions proposed by AAOP[2].

Thus, considering the need for and importance of epidemiological researches to screen for TMD among Brazilian population, especially in adolescents, and the importance of the AAOP questions in the world context of orofacial pain, the aim of the present study was to test the reliability and the validity of the published Portuguese version of the AAOP questions, using the RDC/TMD Axis I as a reference standard. Our hypothesis is that a great correspondence between positive answers to the AAOP questions and TMD diagnosis will be found.

MATERIAL AND METHODS

Sampling procedures

The present study is part of a study conducted to explore and to characterize TMD among young Brazilian adolescents. A large sample was selected from the overall 12-14 years-old public schoolchildren of Araraquara city, São Paulo state, Brazil (24 schools, n=7,172). The sample size (n=1,257) was based on 2-5% of TMD pain prevalence in adolescents found in previous researches[5,20,22]. This statistical planning assured that we would be able to recruit a sufficient number of participants with TMD. A representative sample was proportionally estimated for each school, based on the number of 12-14-year-old schoolchildren *per* school. Classrooms were selected by chance, and in those selected, students received a brief explanation about TMD and the aims of the research, as a requirement of the Research Ethics Committee. Adolescents were then invited to participate, and an informed consent form was sent to their parents or legal guardian for signature. The letter provided clear instructions about the methodology and the research objectives. Moreover, parents received a sociodemographic questionnaire and an education brochure presenting information about TMD and forms of control. Exclusion criteria consisted of the presence of odontogenic toothache, not having a signed informed consent form and/or the adolescent not agreeing to participate.

Ethical considerations

This study received the full approval of the Research Ethics Committee of Araraquara Dental School, UNESP – Univ. Estadual Paulista (Process #70/10).

The legal guardians of all adolescents who agreed to participate had to sign the informed consent form. Only those who returned the signed informed consent form and stated their agreement to participate were recruited for the evaluations. TMD experts involved in the face and content validity tests also stated their agreement to participate in the study. All the participants were examined at their own schools.

Instruments

Portuguese version[1] of the questionnaire proposed by AAOP for TMD screening[2], which is composed of ten questions with dichotomous answers (yes or no);

Portuguese version[24] of the RDC/TMD[6] Axis I diagnosis, in addition to questions #3 ("Have you had pain in the face, jaw, temple, in front of the ear or in your ear in the past month?"), #4 ("How long ago did your facial pain begin for the first time?") and #14 ("Have you ever had your jaw lock or catch so that it would not open all the way") of Axis II history questionnaire.

Measures

a) Face and content validity of AAOP questions

To keep the semantic, idiomatic, cultural and conceptual equivalence of the Portuguese version, face and content validity processes required 20 dentists, experts in TMD and orofacial pain. To evaluate the comprehension index (CI), the questions were pre-tested in 20 adolescent volunteers from public schools, with the purpose of determining the minimal vocabulary level and clarity of questions. If pertinent, item reformulation was performed until a CI higher than 80% was obtained for the content validity ratio (CVR)[14].

b) Test-retest reliability of AAOP questions and intra-examiner reproducibility of RDC/TMD Axis I

To estimate these characteristics, each instrument was applied to the same adolescents (n=77) from

one school, on two separate occasions, by the same examiner. The time between evaluations was seven days, as used by Campos, et al.[3] (2009). Kappa statistics (k) were used and as agreement score, Landis and Koch[13] (1977) patterns were taken as reference.

c) Reliability and validity of AAOP questions

Two researchers were responsible for visiting the 24 public schools and applying AAOP questions and RDC/TMD Axis I to the adolescents. One trained examiner first applied the AAOP questions. Then a second trained examiner, blinded to the results of the AAOP questions, performed the RDC/TMD physical examinations.

The reliability of the overall scale and each item was estimated by internal consistency, using the Kuder-Richardson coefficient (KR-20). Inter-item correlation values were also computed, using Pearson correlation coefficients. Results were compared with adequate values, i.e. KR-20≥0.70 and an inter-item correlation higher than 0.20.

After obtaining a reliable scale, the validity study was conducted using RDC/TMD Axis I as reference standard. According to the results, individuals were classified according to Axis I diagnostic combinations:

No TMD: no myofascial pain, disk displacements or arthralgia/osteoarthritis;

Overall TMD: myofascial pain, disk displacements and arthralgia/osteoarthritis, isolated or combined;

Painful TMD: myofascial pain and/or arthralgia/osteoarthritis.

d) Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and receiver operating characteristic (ROC curve)

To define which questionnaire items were mainly associated with the diagnosis of TMD overall and painful TMD, multivariate logistic regression models (for dichotomous variables) were used. The models consisted of a stepwise selection of variables, identifying the most important items related to each

Table 1- Face and content validity results for the American Academy of Orofacial Pain (AAOP) questions (n=20 adolescents and n=20 specialists)

item #	Original AAOP Questions (English)	CI* first	Adapted AAOP Questions (Portuguese)	CI* second	CVR**
1	Do you have difficulty, pain, or both when opening your mouth, for instance, when yawning?	100%	Você tem dificuldade, dor ou ambos ao abrir a boca, por exemplo, ao bocejar?	100%	0.8
2	Does your jaw “get stuck”, “locked” or “go out”?	30%	A sua boca já ficou "travada" ou seu queixo já "caiu"?	100%	0.2
3	Do you have difficulty, pain, or both when chewing, talking, or using your jaws?	90%	Você tem dificuldade, dor ou ambos ao mastigar, falar ou movimentar a boca?	95%	0.7
4	Are you aware of noises in the jaw joints?	45%	Você nota algum barulho perto dos ouvidos quando abre a boca ou mastiga?	100%	0.3
5	Do your jaws regularly feel stiff, tight, or tired?	80%	Normalmente você sente seu rosto cansado, duro ou tenso?	100%	0.1
6	Do you have pain in or near the ears, temples or cheeks?	55%	Você tem dor perto do ouvido, dos lados da cabeça ou nas bochechas?	100%	0.4
7	Do you have frequent headaches, neck aches, or toothaches?	100%	Você tem dores de cabeça, pescoço ou dor nos dentes com frequência?	100%	0.2
8	Have you had a recent injury in your head, neck or jaws?	95%	Recentemente você sofreu alguma pancada na cabeça, pescoço ou queixo?	100%	-0.3
9	Have you been aware of any recent changes in your bite?	100%	Você observou alguma alteração recente na sua mordida sem ter ido ao dentista?	100%	-0.4
10	Have you been previously treated for unexplained facial pain or a jaw joint problem?	75%	Você já recebeu algum tratamento prévio para dor no rosto ou para outro problema na região do ouvido?	100%	-0.7

*CI=Comprehension Index; **Content Validity Ratio considering $CVR_{20,0.05}$=0.42 as critical value

of the mentioned TMD diagnoses.

After conclusion of the selection, we tested the reliability and validity of the residual questions by the same processes described in the previous steps, in order to compare the performance of the reduced items with the original scale. The temporal stability of the final versions was also tested by reproducibility analysis (kappa).

RESULTS

Out of a total of 3,117 adolescents invited to participate in the present study, 1,307 met the inclusion criteria and were present when evaluations were performed (response rate of 41.93%). The majority of the participants were white (67.3%; n=880) and girls (56.8%; n=742). The mean age of the sample was 12.72 years. Mulattos, Afro-Brazilians and Asian participants accounted for 26.8% (n=350), 5.6% (n=73) and 0.3% (n=4), respectively, of the total sample. The diagnoses obtained by RDC/TMD Axis I preseted a total of 397 [30.4% (32.7% F/27.3% M)] adolescents diagnosed positively for any TMD subtype (overall TMD). But, when considering the presence of pain, 330 [25.2% (27.6% F/22.2% M)] presented painful TMD.

a) Face and content validity of AAOP questions

The suggestions of both (adolescents and specialists) were put to the vote and those items first presenting more than 20% of incomprehensibility (#2, #4, #6, and #10) were reformulated (Table 1). Questions were then reapplied and 100% comprehension was obtained for all questions, except question #3 (95%).

Items CVR varied from -0.70 to 0.80. According to the specialists' opinion and considering $CVR_{20,0.05}$=0.42 as critical value, only questions #1 and #3 were considered essential for a TMD investigation.

b) Test-retest reliability of AAOP questions and intra-examiner reproducibility of RDC/TMD Axis I

The test-retest reliability for the AAOP questions was: for questions #1, #4 and #5, moderate agreement (respectively k=0.539, 0.430 e 0.492); for questions #3, #7 and #9, substantial agreement (respectively k=0.642, 0.712 e 0.794); and, for questions #2, #6, #8 and #10, perfect and almost perfect agreements (respectively k=1.0, 0.864, 0.821 and 1.0).

The intra-examiner reproducibility RDC/TMD Axis I was: for myofascial pain, almost perfect agreement (k=0.884); for disk displacements, moderate agreement (k=0.529); and for arthralgia/osteoarthritis, substantial agreement (k=0.795). Considering Axis I diagnostic combinations, for overall TMD substantial agreement was obtained between evaluations (k=0.727), whereas for painful TMD the agreement was considered almost perfect (k=0.856).

c) Reliability and validity of AAOP questions

Table 2 shows the resulting KR-20=0.662 for the ten-item scale. In this analysis, questions #8 and #10 presented a KR-20 higher than the overall scale (respectively 0.673 and 0.668). The same items also demonstrated low and non-significant inter-item correlations, corroborating their low contribution to the questionnaire. Since these items also presented low CVR, they were excluded from the initial questionnaire and the KR-20 was recalculated. The internal consistency of the revised scale was improved to 0.684, being considered

Table 2- Internal consistency results for the overall scale and for the selected items according to the multivariate analysis

Scale items	Overall scale - ten items	Overall scale - eight items	Reduced scale - four items
item #	KR-20*	KR-20*	KR-20*
1	0.606	0.628	-
2	0.655	0.682	-
3	0.610	0.634	-
4	0.632	0.660	0.572
5	0.642	0.670	-
6	0.584	0.609	0.588
7	0.637	0.666	0.667
8	**0.673**	**removed**	**-**
9	0.644	0.672	0.595
10	**0.668**	**removed**	**-**
SCALE	0.662	**0.684**	**0.673**

*KR-20: Kurder-Richardson coefficient

adequate (approximately 0.70).

d) Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and ROC curve

After the exclusion of questions #8 and #10, the multivariate regression models showed a significant correlation of items #4, #6, #7 and #9 with both overall TMD and painful TMD diagnosis (Table 3). The internal consistency of those questions was tested again, and the results showed a KR-20 value

Table 3- Results of the multivariate logistic regression model for selecting items, both for overall temporomandibular disorders (TMD) and painful TMD

Overall TMD		Painful TMD	
item #	Summary of stepwise selection*	item #	Summary of stepwise selection*
1	Not significant	1	Not significant
2	Not significant	2	Not significant
3	Not significant	3	Not significant
4	p=0.0001	4	p=0.003
5	Not significant	5	Not significant
6	p<0.0001	6	p<0.0001
7	p<0.0001	7	p<0.0001
8	-	8	-
9	p<0.0001	9	p<0.0001
10	-	10	-

*Significance of chi-square tests

Table 4- Validity results according to number of positive responses to the eight-item and four-item questionnaires, for the diagnosis of overall temporomandibular disorders (TMD) and painful TMD

Scale	Diagnosis	Cut-off value	Sensitivity	Specificity	PPV	NPV	Accuracy	Sum: Sensitivity + PPV
8	Overall	>0	94.2	38.1	39.9	93.8	55.2	134.1
		>1	84.4	71.4	56.3	91.3	75.2	140.7
		>2	**67.5**	**90.0**	**74.7**	**86.4**	**83.2**	**142.2**
		>3	44.6	96.6	85.1	80.0	80.8	129.7
		>4	25.9	99.3	94.5	75.5	77.1	120.4
		>5	13.6	99.9	98.2	72.6	73.6	111.8
		>6	6.1	100.0	100.0	70.9	71.5	106.1
	Painful TMD	>0	98.8	37.5	34.5	98.9	52.9	133.3
		>1	94.6	71.0	52.4	97.5	76.9	147.0
		>2	**78.2**	**89.7**	**71.9**	**92.4**	**86.7**	**150.1**
		>3	51.5	96.1	81.7	85.4	84.8	133.2
		>4	30.6	99.2	92.7	80.9	81.8	123.3
		>5	16.4	99.9	98.2	78.0	78.8	114.6
		>6	7.3	100.0	100.0	76.2	76.7	107.3
4	Overall	>0	90.4	65.5	53.3	94.0	73.1	143.7
		>1	**64.2**	**91.4**	**76.6**	**85.4**	**83.2**	**140.8**
		>2	35.5	98.1	89.2	77.7	79.1	124.7
		>3	16.1	99.9	98.5	73.2	74.6	114.6
	Painful TMD	>0	96.7	63.8	47.4	98.3	72.1	144.1
		>1	**73.9**	**90.9**	**73.3**	**91.2**	**86.6**	**147.2**
		>2	40.9	97.7	85.4	83.0	83.3	126.3
		>3	19.1	99.8	96.9	78.5	80.0	116.0

TMD: temporomandibular disorders; PPV: positive predictive values; NPV: negative predictive values

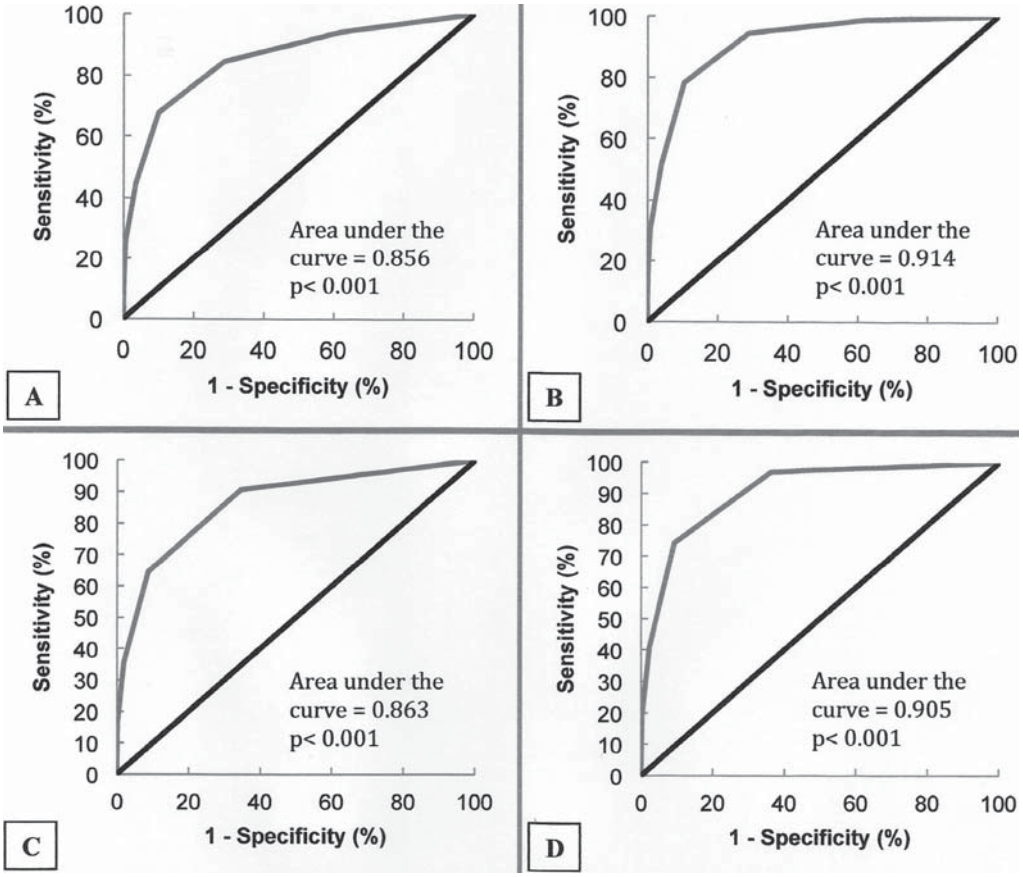

Figure 1- Receiver operating characteristic (ROC) curves for the diagnosis of overall temporomandibular disorders (TMD) and painful TMD with the eight-item and the four-item questionnaires, compared with the reference standard [Research Diagnostic Criteria for Temporomandibular Disorders (RDC/TMD) Axis I]. (A) ROC curve for diagnosing overall TMD with the eight-item questionnaire; (B) ROC curve for diagnosing painful TMD with the eight-item questionnaire; (C) ROC curve for diagnosing overall TMD with the four-item questionnaire; and (D) ROC curve for diagnosing painful TMD with the four-item questionnaire

quite similar to that obtained with the eight-item scale (0.673) (Table 2). Inter-item correlations presented for the final versions were all significant and adequate.

Using the reliable scales, the validity of the eight-item and the four-item questionnaires was tested considering the results obtained with the reference standard. Similar good results were obtained by using both short and long versions. As can be observed in Table 4, the best thresholds for detecting overall TMD and painful TMD were in more than two positive answers (>2) for the eight-item questionnaire, and in more than one positive answers (>1) for the four-item version, since they provided the best balance between sensitivity and specificity in ROC curve analysis. For both versions, excellent specificity was demonstrated (about 90%); however, the questions showed better ability in correctly detecting individuals with painful TMD in comparison with overall TMD. Moreover, the best accuracy values were obtained at the same cut-off values, i.e., at the same number of positive responses (Table 4). Referred ROC curves are presented in Figure 1.

The temporal stability of the final versions was tested in the sample (n=77) again. For the eight-item version, in the cut-off >2 positive answers, kappa value was 0.655, whereas for the four-item version, in the cut-off >1 positive answers, k=0.840.

DISCUSSION

The most important finding of the present study was that the adapted Portuguese version of the AAOP questions showed both good reliability and validity for screening TMD in Brazilian adolescents. Basically, the goal of obtaining a validated screener is to use it in a cost- and time-effective manner[12]. Screening tests can be applied in large samples, such as in epidemiological studies, to delineate those individuals who need further evaluation[7]. Screeners are also important for primary care providers to detect a range of conditions and manage the problem or refer it to specialists when necessary. Thus, for the diagnosis of diseases based

on symptoms (e.g., TMD, migraine), questionnaire-based screening is an interesting approach, because it is safe and relatively inexpensive to use[16].

As previously mentioned, the majority of published TMD screeners[3,12,21,29,30] present many items overlapping with the AAOP questions for the same purpose. Screening tools are validated by using a diagnostic reference standard (herein, the RDC/TMD Axis I). The term "reference standard" does not imply that the diagnosis is error-free, but that it was the best available diagnostic criterion at the time of the study[16]. The RDC/TMD are the most important diagnostic tool, properly translated into Portuguese[24] and other languages, in addition to being adapted and validated, and have been extensively used since 1992[6]. Today, revised criteria, known as Diagnostic Criteria for Temporomandibular Disorders (DC/TMD), have been presented[25], but the older RDC/TMD continue the most acceptable and well-known standard for diagnosing TMD in researches, and that is the reason they selected as reference standard in this study, as it has previously been done by others[4,21,30]. According to the literature, the RDC/TMD Axis I can be applied to both adults and adolescents[17,29].

The first step in testing the questions is to assure that adolescents will easily be able to respond to them. Similar to a previous study[29], our results demonstrated that participants were able to completely understand the questions, which is of particular importance, since they were not specifically designed to be applied in adolescents. The understanding of the questions' meaning is an essential part and, in a recent research[4], authors attributed the low sensitivity/specificity of the AAOP TMD screening questionnaire in a sample of adolescents to the language used in the self-explanatory questions and to difficulties in comprehension.

The reliability analysis of the questions is another extremely important stage and it can be achieved by calculating the reproducibility and the inter-item correlation, as well as by the internal consistency of the instrument[3]. For the test-retest reliability and reproducibility study, the literature suggests an interval of up to one month between the two applications. However, for TMD investigation the interval between the measures should be reduced, due to the intermittent characteristics of its signs and symptoms. As it has previously been recommended[3], an interval of only one week was chosen for this study. Although an acceptable temporal stability[13] was obtained for all questionnaire items and for both short and long versions in their respective optimal cut-off points, the inter-item correlation analysis showed that items #8 and #10 were not significantly correlated with each other. In addition, after removing those items from the questionnaire, the internal consistency value improved – being closer to 70.0%, considered adequate. Accordingly, questions #8 and #10 – which intend to evaluate the history of recent injuries in the head, neck and jaws and the history of any previous treatment for facial pain or jaw problems, respectively – do not seem to be good predictors of a TMD diagnosis. Hence, the eight-item questionnaire presented better ability to detect TMD in the studied population. This is in agreement with other published screeners, questionnaires and checklists, in which injuries in the head, neck and jaws and the history of any previous treatment for facial pain or jaw problems are not explored[10,12,21,29,30].

The short version of the questionnaire, with only four items, also showed good reliability presenting a KR-20 slightly lower than that achieved with the long version, as well as significant and adequate inter-item correlations. Although the results of the multivariate analysis contradict the opinion of experts, who considered questions #1 and #3 the most important items of the questionnaire, the residual items show agreement with a questionnaire previously used for screening TMD in a Brazilian adult population[11]. Question #4, as it regards the self-report of noises in the jaw joints, is present in old[10,29] and recently[30] published screeners. Question #6 is part of question #3 of the RDC/TMD[7] Axis II used as the reference standard for a positive painful TMD diagnosis, which also overlaps with other recently-developed screeners[12,21,30]. Question #7 is also presented in other screeners[29,30] and this is justified because of the strong associations reported especially between headaches[9,15,19] and neck aches[28] with TMD. Finally, question #9, about occlusal changes, probably evolved from an older proposed screener[10] and makes sense given that a recent study has demonstrated that perceived dental changes were among the most common perceived symptoms in adolescents with TMD[19].

It is also important to measure the predictive validity of a questionnaire. Although perfect sensitivity and specificity would be ideal, in clinical practice, the optimal balance depends upon the consequences of classification and misclassification[16]. Once TMD, in general, is not a severe condition and a certain degree of misclassification is acceptable, it is suggested that a clinically useful TMD investigation test needs to present a minimum of 70% sensitivity and 75-95% specificity[6,10]. Thus, we defined the best cut-off points for both diagnoses as more than two positive responses (>2) for the long version and more than one positive response (>1) for the short version because they provided the best balance between sensitivity and specificity values, as previously recommended[12]. Our results are in

agreement with a recent study in which, by using the AAOP questionnaire among adolescents, the values of specificity were observed to be higher than sensitivity values[4]. However, maybe due to the differences in the methodology employed, we were able to reach higher values of sensitivity and specificity. The ROC curve analysis also confirmed good performances of the questionnaire as a function of sensitivity and specificity, once the areas presented under the curve were above 85% in all the four graphics (Figure 1).

There was a high prevalence of TMD pain observed in our sample (25.2%), in comparison to the previous literature reports (of 2-5%)[5,20,22]. Nevertheless, this variability is not surprising, once the TMD prevalence reported in children and adolescents tends to vary considerably within literature due to methodological differences[27]. This can be also attributed to, at least, two more reasons. As previously commented in our recent publication[9], the prevalence of TMD might be higher in developing countries, since low social status has comprised a risk factor for TMD pain in adults according to a recent study conducted in Brazil[18]. Moreover, the sample invitation procedure consisted of a brief explanation about TMD and its associated symptoms (e.g., headaches, joint sounds, facial pain, limitation of movements) as a requirement of the Research Ethics Committee, and it is possible that adolescents were interested in participating because they had previously noticed some of the symptoms cited by the researchers during the invitation procedure. This might have caused a selection bias that must be considered for the high prevalence of TMD observed in the present study.

An efficient screening test requires a high PPV, but PPVs obtained in the present study are not universally applicable. PPV is determined by 3 factors: the sensitivity of the test, the specificity of the test, and the prevalence of the condition in the population being tested[16]. The higher the prevalence of individuals with a particular condition, the higher the probability of a test to detect the condition, thus increasing the PPV of the test. This is the reason why PPVs for diagnosing overall TMD were higher than for painful TMD, in both short and long versions. In fact, there are people who do not have pain but have received a TMD diagnosis according to RDC/TMD (e.g., disc displacements without pain), leading to a higher prevalence of overall TMD. Furthermore, it is assumed that a screening test is excellent when the sum of the sensitivity and PPV (in percentage) equals at least 150[7]. Thus, the high PPVs obtained due to the high prevalence of the condition in the sample may have influenced the final scores presented.

The results presented show that in both versions the psychometrical properties evaluated remained almost unaltered. The applicability of the short or long version will depend on the clinician's aims and on the resources available to conduct the investigation. But, as proposed, AAOP questions accomplish a good screening for TMD, especially for painful TMD. However, improvements should be made before the questionnaire can be used efficiently, and we encourage further research to ensure adequate psychometric characteristics. Moreover, a better-designed TMD screener questionnaire has been proposed recently for adults[12] and its use should be considered for further research on screening TMD.

Indeed, some limitations of the present study deserve attention. Firstly, no translation and cultural equivalence has been conducted from the original English version of the questions[2], which is recommended[23]. We assumed that since these questions were gathered from the Portuguese[1] version of the AAOP guidelines, they have probably resulted from a consensus among the book translators. It should be highlighted that, when a non-English scale version is tested, the original one (in English) should be presented within the publication, accompanied by the English retro-translation, when conducted[8]. Thus, we emphasize that AAOP items should be translated and validated as recommended before testing their psychometric properties in other languages (non-Brazilian Portuguese).

Secondly, we were not able to measure the TMD graded chronic pain scale because the RDC/TMD Axis II is not properly validated in English or in Portuguese for use in adolescents. Certainly, it will be possible to assess TMD severity and other features in future evaluations, after the screening test[7]. Thirdly, it is important to highlight that the psychometric properties observed here are related to the use of the aforementioned questions in our sample, which might have been influenced by the high prevalence of TMD observed in this sample. Lastly, we did not make any differential diagnosis other than the RDC/TMD Axis I, as others have done[12,30].

The strengths of this study include the use of standardized methodology for TMD classification together with a well-established method to assess the reliability and validity of measures. In addition, the number of participants in the present study was much higher than it was in other similar researches[4,21,29].

CONCLUSION

In conclusion, we found great agreement between the AAOP screening questionnaire and a positive TMD diagnosis, especially for painful TMD, according to the RDC/TMD Axis I. Nevertheless, the

use of a screener should never substitute the "gold standard" for a clinical diagnosis, which consists of combining patients' self-reported symptoms with a confirmatory clinical examination[2,10]. We recommend the use of the questions when the TMD screening is intended for application in young Brazilian adolescents.

ACKNOWLEDGMENTS

The authors thank São Paulo Research Foundation (FAPESP) for the financial assistance (Process #2010/20445-9), Fernanda Salloume Sampaio Bonafé for the databasis, and Creusa Maria Roveri Dal Bó for the statistical analysis.

REFERENCES

1- Academia Americana de Dor Orofacial. Avaliação geral do paciente com dor orofacial. In: De Leeuw R, editor. Dor orofacial: guia de avaliação, diagnóstico e tratamento. 4th ed. São Paulo: Quintessence; 2009. p. 25-47.

2- American Academy of Orofacial Pain. General assessment of the orofacial pain patient. In: De Leeuw R, Klasser GD, editors. Orofacial pain - guidelines for assessment, diagnosis and management. 5th ed. Chicago: Quintessence; 2013. p. 25-46.

3- Campos JA, Gonçalves DA, Camparis CM, Speciali JG. Reliability of a questionnaire for diagnosing the severity of temporomandibular disorder. Rev Bras Fisioter. 2009;13(1):38-43.

4- De Santis TO, Motta LJ, Biasotto-Gonzalez DA, Mesquita-Ferrari RA, Fernandes KP, Godoy CHL, et al. Accuracy study of the main screening tools for temporomandibular disorder in children and adolescents. J Bodyw Mov Ther. 2014;18(1):87-91.

5- Drangsholt M, LeResche L. Temporomandibular disorder pain. In: Crombie IK, Croft PR, Linton SJ, LeResche L, Von Korff M, editors. Epidemiology of pain. Seattle: IASP Press; 1999. p. 203-33.

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

7- Evans MI, Galen RS, Britt DW. Principles of screening. Semin Perinatol. 2005;29(6):364-6.

8- Fournier-Charrière E, Swaine-Verdier A, Tourniaire B, Falissard B. Translation and validation of a behavioural pain scale: what matters most? Pain. 2013;154(6):960-1.

9- Franco AL, Fernandes G, Gonçalves DA, Bonafé FS, Camparis CM. Headache associated with temporomandibular disorders among young Brazilian adolescents. Clin J Pain. 2014;30(4):340-5.

10- Gerstner GE, Clark GT, Goulet JP. Validity of a brief questionnaire in screening asymptomatic subjects from subjects with tension-type headaches or temporomandibular disorders. Community Dent Oral Epidemiol. 1994;22(4):235-42.

11- Gonçalves DA, Dal Fabbro AL, Campos JA, Bigal ME, Speciali JG. Symptoms of temporomandibular disorders in the population: an epidemiological study. J Orofac Pain. 2010;24(3):270-8.

12- Gonzalez YM, Schiffman E, Sharon M, Seago B, Truelove EL, Slade G, et al. Development of a brief and effective temporomandidibular disorder pain screening questionnaire: reliability and validity. J Am Dent Assoc. 2011;142(10):1183-91.

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

14- Lawshe CH. A quantitative approach to content validity. Pers Psychol. 1975;28(4):563-75.

15- LeResche L, Mancl LA, Drangsholt MT, Huang G, Von Korff M. Predictors of onset of facial pain and temporomandibular disorders in early adolescence. Pain. 2007;129(3):269-78.

16- Lipton RB, Bigal ME, Amatniek JC, Stewart WF. Tools for diagnosing migraine and measuring its severity. Headache. 2004;44(5):387-98.

17- List T, Dworkin SF. Comparing TMD diagnoses and clinical findings at Swedish and US TMD centers using research diagnostic criteria for temporomandibular disorders. J Orofac Pain. 1996;10(3):240-53.

18- Magalhães B, de-Sousa S, Mello V, da-Silva-Barbosa A, de-Assis-Morais M, Barbosa-Vasconcelos M, et al. Risk factors for temporomandibular disorder: binary logistic regression analysis. Med Oral Patol Oral Cir Bucal. 2013 Dec 7. [Epub ahead of print].

19- Moyaho-Bernal A, Lara-Muñoz MD, Espinosa-De Santillana I, Etchegoyen G. Prevalence of signs and symptoms of temporomandibular disorders in children in the State of Puebla, Mexico, evaluated with the research diagnostic criteria for temporomandibular disorders (RDC/TMD). Acta Odontol Latinoam. 2010;23(3):228-33.

20- Nilsson IM, List T, Drangsholt M. Prevalence of temporomandibular pain and subsequent dental treatment in Swedish adolescents. J Orofac Pain. 2005;19(2):144-50.

21- 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.

22- Nilsson IM, List T, Drangsholt M. Incidence and temporal patterns of temporomandibular disorder pain among Swedish adolescents. J Orofac Pain. 2007;21(2):127-32.

23- Ohrbach R, Bjorner J, Jezewski M, John MT, Lobbezoo F. Guidelines for establishing cultural equivalency of instruments. [online]. 2009. [cited February 23, 2014] Available from: http://www.cost-td1005.net/fileadmin/user_upload/Dokumente/Guidelines_for_Translation_and_Cultural_Equivalency_of_Instruments_-_2013_05_11.pdf.

24- Pereira Júnior FJ, Favilla EE, Dworkin S, Huggins K. Research diagnostic criteria for temporomandibular disorders (RDC/TMD): formal translation to portuguese. JBC J Bras Clin Odontol Integr. 2004;8(47):384-95.

25- Schiffman E, Ohrbach R, Truelove E, Look J, Anderson G, Goulet J, et al. Diagnostic Criteria for Temporomandibular Disorders (DC/TMD) for Clinical and Research Applications: recommendations of the International RDC/TMD Consortium Network and Orofacial Pain Special Interest Group. J Oral Facial Pain Headache. 2014;28(1):6-27.

26- Schiffman EL, Truelove EL, Ohrbach R, Anderson GC, John MT, List T, et al. The Research Diagnostic Criteria for Temporomandibular Disorders. I: overview and methodology for assessment of validity. J Orofac Pain. 2010;24(1):7-24.

27- Sena MF, Mesquita KS, Santos FR, Silva FW, Serrano KV. Prevalence of temporomandibular dysfunction in children and adolescents. Rev Paul Pediatr. 2013;31(4):538-45.

28- Vierola A, Suominen AL, Ikavalko T, Lintu N, Lindi V, Lakka HM, et al. Clinical signs of temporomandibular disorders and various pain conditions among children 6 to 8 years of age: the PANIC study. J Orofac Pain. 2012;26(1):17-25.

29- Wahlund K, List T, Dworkin SF. Temporomandibular disorders in children and adolescents: reliability of a questionnaire, clinical examination, and diagnosis. J Orofac Pain. 1998;12(1):42-51.

30- Zhao NN, Evans RW, Byth K, Murray GM, Peck CC. Development and validation of a screening checklist for temporomandibular disorders. J Orofac Pain. 2011;25(3):210-22.