

Severity of temporomandibular disorders in women: validity and reliability of the Fonseca Anamnestic Index

Juliana Alvares Duarte Bonini Campos^(a)
Andréa Corrêa Carrascosa^(b)
Fernanda Salloume Sampaio Bonafé^(c)
João Maroco^(d)

^(a)Departamento de Odontologia Social, Faculdade de Odontologia de Araraquara, Universidade Estadual Paulista - Unesp, Araraquara, SP, Brasil.

^(b)Departamento de Alimentos e Nutrição, Faculdade de Ciências Farmacêuticas de Araraquara, Universidade Estadual Paulista - Unesp, Araraquara, SP, Brasil.

^(c)Departamento de Materiais Restauradores e Prótese, Faculdade de Odontologia de Araraquara, Universidade Estadual Paulista - Unesp, Araraquara, SP, Brasil.

^(d)Departamento de Ciências Psicológicas, Instituto Universitário - ISPA, Lisboa, Portugal.

Declaration of Interests: The authors certify that they have no commercial or associative interest that represents a conflict of interest in connection with the manuscript.

Corresponding Author:

Juliana Alvares Duarte Bonini Campos
E-mail: jucampos@foar.unesp.br

<http://dx.doi.org/10.1590/S1806-83242013005000026>
Epub Nov 25, 2013

Submitted: May 03, 2013
Accepted for publication: Jul 31, 2013
Last revision: Sep 30, 2013

Abstract: The aim of this study was to assess the validity and reliability of the Fonseca Anamnestic Index (IAF), used to assess the severity of temporomandibular disorders, applied to Brazilian women. We used a probabilistic sampling design. The participants were 700 women over 18 years of age, living in the city of Araraquara (SP). The IAF questionnaire was applied by telephone interviews. We conducted Confirmatory Factor Analysis (CFA) using Chi-Square Over Degrees of Freedom (χ^2/df), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA) as goodness of fit indices. We calculated the convergent validity, the average variance extracted (AVE) and the composite reliability (CR). Internal consistency was assessed by Cronbach's alpha coefficient (α). The factorial weights of questions 8 and 10 were below the adequate values. Thus, we refined the original model and these questions were excluded. The resulting factorial model showed appropriate goodness of fit to the sample ($\chi^2/df = 3.319$, CFI = 0.978, TLI = 0.967, RMSEA = 0.058). The convergent validity (AVE = 0.513, CR = 0.878) and internal consistency ($\alpha = 0.745$) were adequate. The reduced IAF version showed adequate validity and reliability in a sample of Brazilian women.

Descriptors: Reproducibility of Results; Validity of Tests; Scales; Temporomandibular Joint Disorders.

Introduction

Temporomandibular disorder (TMD) is a collective term embracing all the problems regarding the temporomandibular joint and related musculoskeletal structures.¹

The literature shows wide-ranging variation in TMD prevalence and TMD symptoms in different populations. This may be attributed to the variety of study designs, sampling methods, measurement instruments, and different diagnostic criteria for TMD. In a systematic literature review, Manfredini *et al.*² found a TMD prevalence from 2.6% to 11.4% in the normative population, in different countries. Moreover, Gonçalves *et al.*³ found a 39.2% prevalence of TMD symptoms in the Brazilian population, in which women were significantly more likely to have TMD than men (RR > 1.0; $p < 0.001$).

Although screening for TMD in the population is the concern of researchers, different facets of TMD and different measuring instruments have been used in different studies (e.g. TMD-type diagnosis, symptoms, severity and/or mandibular limitation).

The most common and important instrument in assessing TMD is the Research Diagnostic Criteria for Temporomandibular Disorders - RDC/TMD. RDC/TMD has been used in several clinical and epidemiological studies. It is composed of two evaluation axes. Axis I conducts a clinical evaluation requiring the presence of the patient. Thus, its use in epidemiological studies, where data is gathered by telephone, mail or internet, is difficult or may even be impossible. With this in mind, other instruments have been developed for evaluating TMD. Fonseca's Anamnestic Index (IAF)⁴ has been proposed as a low cost and easy to apply alternative and has been used in screening for TMD in a non-patient population.

In Brazil, Fonseca's Anamnestic Index (IAF) has frequently been used⁵⁻⁸ to classify individuals according to TMD severity (mild, moderate, severe and no TMD), and also to screen patients for further developments in diagnosing TMD.⁹ The IAF is a scale proposed in the Portuguese language, consisting of 10 questions whose answers are arranged in a three-point scale format (no, sometimes, yes). Despite its frequent use in Brazil, the IAF has not been applied in other countries. However, the simplicity of its application, and its dispensing with the requirement of a physical examination of the patient, makes it suitable for fast epidemiological screening by telephone, mail or internet surveys.

Measurement theory states that the use of psychometric and/or clinimetric scales, addressing latent variables, requires the assessment of the metric qualities of the data gathered with these scales in the study sample, so that the researcher may be confident about the validity and reliability of the data gathered. Despite the widespread use of the IAF, we did not find studies in the literature that addressed its validity to measure the construct of "severity of temporomandibular disorders." Moreover, studies of IAF's metric properties have been limited to the evaluation of its reliability¹⁰ and predictive validity.⁴

Thus, our study was conducted to evaluate the

validity (construct and concurrent), and reliability of Fonseca's Anamnestic Index (IAF), applied to Brazilian women.

Methodology

Study and sample design

This is a cross-sectional validation study. The sampling design was probabilistic, and conducted in two stages. In the first stage, we stratified the participants, according to their census segment. In the second stage, we randomly and systematically selected each participant from the phonebook.

The participants lived in the city of Araraquara (SP, Brazil), were female, and over 18 years of age.

The sample size was established from an expected prevalence of TMD symptoms of approximately 40% in the female population, as presented by Gonçalves *et al.*³ We estimated the population of adult women in Araraquara municipality at about 70,000. The significance level was 5%, the test power was 80%, and the sampling error was set at 10%. Assuming a non-response rate of 20%, we estimated a minimum sample size of 715.

Instrument

The instrument used was the questionnaire proposed in Brazilian Portuguese by Fonseca *et al.*⁴ to estimate Fonseca's Anamnestic Index (IAF). The IAF is a one-dimension questionnaire, consisting of 10 questions with a three-point scale (0 = no, 5 = sometimes and 10 = yes). The IAF has been widely used in Brazilian studies to estimate the severity of temporomandibular disorders.

Procedures

Supported by evidence from the literature and considering the simplicity of the instrument's application, we chose to conduct the data collection through telephone interviews, not to exceed 5 minutes.

When the respondents answered the call, the researcher identified herself and read the free and informed consent statement for the research to be undertaken. Only those subjects who agreed to its terms participated in the study. The calls were always made by the same interviewer, calibrated in a pilot study. The intra-reproducibility of the re-

searcher doing the calls, regarding the classification of TMD (mild, moderate, severe and no TMD), was estimated at two distinct time periods, one week apart. The reproducibility was considered good ($n = 62$; $\kappa = 0.89$).

Analysis of the psychometric characteristics

The IAF construct validity was estimated by the factorial and convergent validity. For the purpose of estimating the factorial validity, we used Confirmatory Factor Analysis (CFA), applying the polychoric correlation matrix performed on the MPLUS 6.0 (Muthén & Muthén, Los Angeles, USA). In general, CFA is an appropriate method of evaluation, if the items are good indicators of the proposed factors a priori to the scale, in a given sample. If the original model fits the variance-covariance matrix of the items suitably, it is indicative of factorial validity. Goodness of fit for individual items was evaluated through their factorial weights (λ). Standardized lambda values greater than 0.5 were considered indicative of good fit of the items to the scale. The global goodness of fit was evaluated using the standard goodness of fit indices for CFA:

- Chi-Square over Degrees of Freedom (χ^2/df),
- Comparative Fit Index (CFI),
- Tucker-Lewis Index (TLI), and
- Root Mean Square Error of Approximation (RMSEA).

The fit of the model to the variance and covariance data of the items was considered adequate when $\chi^2/df \leq 4.0$, $CFI \geq 0.90$, $TLI \geq 0.90$ and $RMSEA \leq 0.10$. We refined the model by also taking into consideration the modification indices. Accordingly, the items that presented Lagrange Multipliers (LM) > 11 ($p < 0.001$), suggesting correlation between errors of measurement, were excluded.

The convergent validity was evaluated using the Average Variance Extracted (AVE) and the Composite Reliability (CR), as proposed by Fornell and Larcker.¹¹ These were considered adequate if $AVE \geq 0.50$ and $CR \geq 0.70$.

Internal consistency was estimated using Cronbach's alpha coefficient (α) and was considered ad-

equated if $\alpha \geq 0.70$.¹² Since α underestimates the true reliability for congeneric items, as is the case of the IAF, Composite Reliability (CR) was also evaluated to substantiate the reliability of the IAF in the present sample.

Concurrent validity was assessed by correlation analysis with the mandibular function impairment questionnaire (MFIQ), duly validated for the population under study, according to Campos *et al.*¹³

Ethical aspects

The present study was approved by the Ethics Committee of the School of Pharmaceutical Sciences - UNESP (protocol 52/2009).

Results

The response rate (RR) obtained was 97.9%. A total of 700 women with a mean age of 44.3 (SD = 16.3) years, participated in this study. Of the participants, only 39.0% used dental prosthetics and 55.4% reported taking some type of chronic medication. Regarding marital status, 22.9% were single, 58.1% were married, 11.2% were widowed, and 7.9% were divorced. As for average monthly household income, 1.1% of participants reported having an income of USD 5676.00, 36.9% reported USD 1857.00, 55.3% reported USD 682.50, and 6.7% reported USD 317.00.

The distribution of the participants, according to the answers to each question of Fonseca's Anamnestic Index (IAF) is shown in Table 1.

There was a high prevalence of "No" answers to the IAF questions. This is because the sample was composed of a normative population.

Figure 1 shows the IAF's full factorial model after its refinement, states the standardized factorial weights, and explains the variance for each question, respectively.

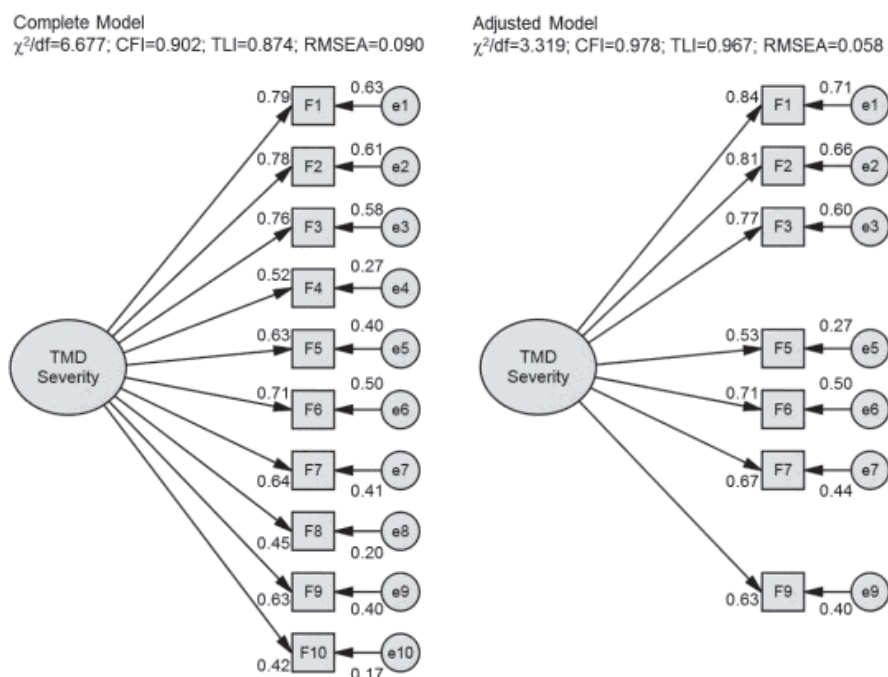
In the complete factorial design, questions 8 and 10 had factorial weights below the recommended values ($\lambda < 0.50$), and question 4 proved to be strongly correlated with other items, according to the modification indices (LM > 11). After these items were removed, there was a better fit of the factorial structure to the data.

The adjusted model showed good convergent

Table 1 - Distribution of participants according to the responses to each question of Fonseca’s Anamnestic Index (IAF).

Question	Answer, n(%)		
	No	Sometimes	Yes
1. Do you have difficulty opening your mouth wide?	612 (87.4)	46 (6.6)	42 (6.0)
2. Do you have difficulty moving your jaw to the sides?	612 (87.4)	44 (6.3)	44 (6.3)
3. Do you feel fatigue or muscle pain when you chew?	526 (75.1)	94 (13.4)	80 (11.4)
4. Do you have headaches?	389 (55.6)	115 (16.4)	196 (28.0)
5. Do you have neck pain or a stiff neck?	408 (58.3)	141 (20.1)	151 (21.6)
6. Do you have ear aches or pain in that area (temporomandibular joint)?	544 (77.7)	72 (10.3)	84 (12.0)
7. Have you ever noticed any noise in your temporomandibular joint while chewing or opening your mouth?	538 (76.9)	62 (8.9)	100 (14.3)
8. Do you have any habits such as clenching or grinding your teeth?	431 (61.6)	56 (8.0)	213 (30.4)
9. Do you feel that your teeth do not come together well?	450 (64.3)	49 (7.0)	201 (28.7)
10. Do you consider yourself a tense (nervous) person?	185 (26.4)	135 (19.3)	380 (54.3)

Figure 1 - Confirmatory Factor Analysis of Fonseca’s Anamnestic Index (IAF), the complete and refined models. (χ^2/df : Chi-Square Over Degrees of Freedom; CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; RMSEA: Root Mean Square Error of Approximation).



validity (AVE = 0.513, CR = 0.878), concurrent validity ($r = 0.66$, $p < 0.01$) and internal consistency ($\alpha = 0.745$).

Discussion

Knowledge of the metric characteristics of the data collected with a given instrument, when applied to different samples and different formats

(telephone, mail, internet and personal interview), is an essential prerequisite for using this data, given that this procedure is the only way to check the quality of the data collected. However, this methodology has been widely neglected in the literature. This situation can be explained through the training given to health professionals,¹⁴⁻¹⁶ who have technical difficulties in determining the validity and reliability

of statistical analyses, or who lack knowledge about the importance of performing this procedure when using scales. It is important to highlight that the metric properties of the data, collected with a given instrument, depend on the characteristics of the sample;¹³ therefore, those properties should always be tested and presented along with the study results.

RDC/TMD is one of the most widely used instruments in TMD diagnosis. However, one of its dimensions (Axis I) requires the presence of the patient, thus limiting the use of RDC/TMD for some studies. Conversely, the IAF is a questionnaire composed of 10 items that may be answered by the patient by telephone, mail, or internet surveys. When considering large initial epidemiological studies, this feature of the IAF represents a great advantage for quick-to-apply, cost-effective and large scale epidemiological studies.

Fonseca's Anamnestic Index (IAF) is a scale proposed to measure the "severity of temporomandibular disorders" construct. Therefore, it was built using a three-point ordinal scale, which calls for a polychoric correlation matrix to estimate the model's fit to the sample. However, we should stress that, for this analytical model to have adequate statistical power, a large sample size must be used;¹⁷ on the other hand, this can be a limiting factor for testing the metric characteristics when using IAF in clinical samples. Nevertheless, this limitation does not invalidate the professional's responsibility of providing data on the reliability and validity of the IAF, when using the index in different contexts.

The original IAF structure, when applied to normative women, was not adequate; the fit of the model to the sample had to be improved (Figure 1); this implied the removal of three items. The inadequacy of these questions had already been reported by Bevilaqua-Grossi *et al.*⁵ and by Campos *et al.*,¹⁰ who reported the low internal consistency and low predic-

tive ability of these items. The contribution of these items to the "severity of temporomandibular disorders" construct had not yet been assessed. Thus, it was difficult to establish a direct comparison between the results gathered in this study and other studies that have used the IAF. However, it is important to note, along with the results shown in Figure 1, that the different metric characteristics observed in the studies mentioned above point to the inadequacy of questions 4, 8 and 10, and reinforce the need to remove them in order to obtain more reliable information and valid diagnosis of TMD severity. The lack of fit of these 3 questions to the IAF may account for the fact that these questions do not evaluate the structural-anatomic alterations related to the temporomandibular joint function, as the other questions do.

After the number of items was reduced, adequate factorial and convergent validity and internal consistency of the IAF was observed. The reduced scale was, therefore, able to capture the "severity of temporomandibular disorders" construct with higher psychometric sensitivity, validity and reliability.

It must be pointed out that the simpler IAF version, proposed in this study, requires further external validation. Since the items are ordinal with 3 points, requiring that an analysis of polychoric correlations be made, future studies will require large sample sizes. The lack of external validation of the proposed refined IAF is a limitation, insofar as only females participated in this study. However, data gathered here is a first contribution to the study of the psychometric properties of data gathered with the IAF, in terms of both construct-related validity and reliability for the probabilistic sample used.

Conclusion

A reduced version of Fonseca's Anamnestic Index (IAF) showed adequate validity and reliability for the sample of women studied.

References

1. American Society of Temporomandibular Joint Surgeons. Guidelines for diagnosis and management of disorders involving the temporomandibular joint and related musculoskeletal structures. *Cranio*. 2003 Jan;21(1):68-76.
2. Manfredini D, Guarda-Nardini L, Winocur E, Piccotti F, Ahlberg J, Lobbezoo F. Research diagnostic criteria for temporomandibular disorders: a systematic review of axis I epi-

- demologic findings. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2011 Oct;112(4):453-62.
3. 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 Summer;24(3):270-8.
 4. Fonseca DM, Bonfante G, Valle AL, De Freitas SFT. [Diagnóstico pela anamnese da disfunção craniomandibular]. *Rev Gaucha Odontol.* 1994 Jan-Feb;4(1):23-32. Portuguese.
 5. Bevilaqua-Grossi D, Chaves TC, Oliveira AS, Monteiro-Pedro V. Anamnestic index severity and signs and symptoms of TMD. *Cranio.* 2006 Apr;24(2):112-8.
 6. Oliveira AS, Dias EM, Contato RG, Berzin F. Prevalence study of signs and symptoms of temporomandibular disorder in Brazilian college students. *Braz Oral Res.* 2006 Jan-Mar;20(1):3-7.
 7. Nomura K, Vitti M, Oliveira AS, Chaves TC, Semprini M, Siéssere S, et al. Use of the Fonseca's questionnaire to assess the prevalence and severity of temporomandibular disorders in Brazilian dental undergraduates. *Braz Dent J.* 2007;18(2):163-7.
 8. Brandini DA, Pedrini D, Panzarini SR, Benete IM, Trevisan CL. Clinical evaluation of the association of noncarious cervical lesions, parafunctional habits, and TMD diagnosis. *Quintessence Int.* 2012 Mar;43(3):255-62.
 9. Resende CM, Alves AC, Coelho LT, Alchieri JC, Roncalli AG, Barbosa GA. Quality of life and general health in patients with temporomandibular disorders. *Braz Oral Res.* 2013 Apr;27(2):116-21.
 10. Campos JADB, Gonçalves DAG, Camparis CM, Speciali JG. Reliability of a questionnaire for diagnosing the severity of temporomandibular disorder. *Braz J Phys Ther.* 2009 Jan-Feb;13(1):38-43.
 11. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res.* 1981 Feb;18(1):39-50.
 12. Maroco J, Garcia-Marques T. [Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas?]. *Lab Psicol.* 2006;4(1):65-90. Portuguese.
 13. Campos JA, Carrascosa AC, Maroco J. Validity and reliability of the Portuguese version of Mandibular Function Impairment Questionnaire. *J Oral Rehabil.* 2012 May;39(5):377-83.
 14. Altman DG. The scandal of poor medical research. *BMJ.* 1994 Jan;308(6924):283-4.
 15. Schwartz SJ, Sturr M, Goldberg G. Statistical methods in rehabilitation literature: a survey of recent publications. *Arch Phys Med Rehabil.* 1996 May;77(5):497-500.
 16. Kurichi JE, Sonnad SS. Statistical methods in the surgical literature. *J Am Coll Surg.* 2006 Mar;202(3):476-84.
 17. Yung YF, Bentler PM. Bootstrap-corrected ADF test statistics in covariance structure analysis. *Br J Math Stat Psychol.* 1994 May;47(1):63-84.