Community Dent Oral Epidemiol 2005; 33: 125–30

COMMUNITY DENTISTRY AND ORAL EPIDEMIOLOGY

Cross-cultural validation of a short form of the Oral Health Impact Profile for temporomandibular disorders

Marzia Segù[1], Vittorio Collesano[1], Sonia Lobbia[1] and Cristiana Rezzani[2]

Departments of [1]Prosthetic Dentistry and [2]Applied Sanitary Sciences, University of Pavia, Pavia, Italy

Segù M, Collesano V, Lobbia S, Rezzani C. Cross-cultural validation of a short form of the Oral Health Impact Profile for temporomandibular disorders. Community Dent Oral Epidemiol 2005; 33: 125–30. © Blackwell Munksgaard, 2005

Abstract – *Objectives:* We set out to develop and validate an Italian version of the Oral Health Impact Profile Questionnaire (OHIP) that is appropriate for use in temporomandibular disorders (TMD). *Methods:* At first, we had the questionnaire translated from English into Italian by three bilingual individuals whose mother tongue was Italian and thus had three different versions of the questionnaire. These were translated back into English by a native English speaker and the version closest to the original English OHIP was selected. The validation of a questionnaire generally involves the study of the psychometric properties of the instrument: its validity and reliability. Before studying these properties, we assessed the factorial structure of the questionnaire. *Results:* The number of eigenvalues >1, computed by exploratory factor analysis, was seven. The percentage of cumulative variability explained by a model with six dimensions is 66, whereas that explained by a model with seven dimensions is 70. Therefore, considering that the increment of explained variability due to the seventh dimension is low (3.68%) and that the seventh eigenvalue is very close to 1, we considered a six-factor model capable of explaining the factorial structure of the data. Content analysis suggested eliminating the item 'Felt Self-conscious', as most of the subjects did not understand its meaning. Spearman correlation coefficients showed an association between the scores of all the different subscales and the variable for pain. All the coefficients were significantly different from 0 ($P < 0.05$). Cronbach's alpha value, always >0.70, showed quite a good reliability for each of the six subscales. *Conclusions:* These results reveal a reasonable degree of cross-cultural consistency between the two versions of the OHIP, and thus indicate that our Italian version is valid.

Key words: OHIP; validation; TMD

Marzia Segù, Via Battù 79, 27029 Vigevano (PV), Italy
Tel: 39/0381312317
Fax: 39/0381312075
e-mail: marziasegu@interzona.com

Submitted 6 May 2003;
accepted 13 July 2004

Objectives

In recent years, many authors have studied and tested the performance of oral health status measures (1, 2). The Oral Health Impact Profile (OHIP), developed in Australia (3, 4) but now widely in use in other countries, is one these measures. The OHIP is a questionnaire organized into seven sections (functional limitation, physical pain, psychological discomfort, physical disability, psychological disability, social disability and handicap) to study the functional and psychosocial outcomes of oral conditions.

Culture influences people's perception of health, and to compare the oral health-related quality of life of different populations, we first need to verify the cross-cultural equivalence of the instruments used. To date, most cross-cultural studies measuring oral health status have been based on English measures. These were translated into other European

languages (such as Spanish, French and Swedish), using the same method of graduation to obtain the item weights. After this, the weights assigned by lay people and medical professionals were compared. When people from different backgrounds perceive health status in a similar way, we can assume that, in terms of its content, the measure is cross-culturally valid.

Initially, the OHIP questionnaire was adopted in English-speaking countries, and hence no problems with regard to cultural and linguistic equivalence arose. To date, some studies (5–7) have compared the cultural equivalence of the OHIP. Use of the OHIP in oral health studies in Ontario and Québec, Canada, prompted a cross-cultural investigation of perception of health in order to evaluate the equivalence of the measure in the different cultural settings. This evaluation was facilitated by the fact that the OHIP is one of the few oral health-related quality-of-life measures to incorporate item weights.

The aim of our study was to validate an Italian version of the OHIP used by Murray et al. (8). On the basis of the study of Murray et al., we adopted 30 of the 49 items of the original version of the OHIP to minimize respondent burden. We used our version only for patients with temporomandibular disorders (TMD), while other authors made a shortened version of the OHIP appropriate for use in edentulous patients called OHIP-EDENT (9), to introduce a modified short version that is appropriate for TMD patients.

Methods

The questionnaire is divided into two sections, and begins with an introduction that explains the aims of the study, the tasks required, and the confidential nature of all the data collected. Written consent is obtained to indicate that the subject understands the nature and purpose of the proposed study, has had the opportunity to ask questions, and agrees to participate on a voluntary basis.

The first section of the questionnaire collects information on the characteristics of any pain experienced by the subject in the last month. Ten questions investigate the type and location of the pain and associated symptoms, the frequency of the pain, and the approximate duration of a 'pain episode'. Pain severity is measured in two ways: first, by means of a five-point categorization scale that gives the following response options: 'mild', 'uncomfortable', 'moderately severe', 'severe', 'very severe', and secondly, by means of a numerical rating scale, in which 0 indicates 'no pain' and 10 indicates 'pain as bad as it could be'.

The second section of the questionnaire consists of the 30 items of the OHIP itself. Here, the five possible responses are: 'very often', 'fairly often', 'occasionally', 'hardly ever' and 'never'.

Study subjects

We conducted a case–control study. We interviewed 124 new patients treated for TMD pain (mean age 35.1 years; 83.9% women). They were consecutive patients enrolled from the Temporomandibular Disorder Department of the University of Pavia, Italy. A random sample of healthy subjects drawn from the general population ($n = 61$; mean age 41.4 years) was recruited as 'controls'.

All the patients and controls filled in a copy of the questionnaire and then every patient underwent a detailed clinical assessment. All the examinations, consisted of two parts: detailed history analysis with a subjective description of symptoms, and a clinical examination. Both the examinations were conducted by the same researcher.

Back translation

First, we had the questionnaire translated from English into Italian by three bilingual individuals whose mother tongue was Italian. These Italian versions were then translated back into English by another bilingual individual who was a native English speaker. All the translators worked independent of each other. We then compared the three back-translated versions and chose the one most similar to the original English version.

The validation of a questionnaire generally involves investigation of the psychometric properties of the instrument: its validity and reliability. However, validity and reliability studies are possible only in the case of unidimensional scales, i.e. scales that measure only one construct. That is not to say that these properties cannot be investigated in multidimensional scales, but the latter must be broken down to form a series of unidimensional subscales.

Statistical analysis

Factorial structure study

Before studying the psychometric properties of the instrument, we assessed the questionnaire's factorial structure. By means of exploratory factor analysis, we studied the factorial structure of our Italian version of the OHIP: exploratory factor analysis is a multivariate statistical technique that

aims to synthesize a large data set with a smaller number of variables (factors or dimensions), endeavoring to leave out as little as possible.

Exploratory factor analysis indicates the number of dimensions behind the questionnaire items (10). Generally, exploratory factor analysis takes into consideration the number of factors with eigenvalues (extracted by observed correlation matrix) >1. In determining the number of factors, we can consider the percentage of cumulative variability explained by the model.

Validity

The validity of a questionnaire represents the degree to which it measures what it is meant to measure. There are three different types of validity: content validity, criterion-related validity and construct validity.

The purpose of content validation is to assess whether the items truly represent the performance domain or construct of specific interest. Criterion-related validation investigates possible associations between the examined scale and external criteria or other validated measures. As there are no other validated oral-specific health status measures, we took the results of the clinical interview as our external criteria, correlating the OHIP scores with a score corresponding to the sum of the answers to the items investigating pain. As pain was considered as a variable only in the TMD patients (cases), the relevant correlation coefficients were calculated only for the cases. We hypothesized that the greater the pain, the greater is the functional limitation, psychological discomfort, etc. Finally, construct validation studies the theoretical relationships between the items of a scale and a grid of other constructs, possibly associated with the construct measured by the questionnaire. One of the easiest methods of construct validation is that which employs known-groups analysis. It is based on the principle that subjects belonging to different groups will almost certainly respond differently to the questionnaire. If the questionnaire is valid, it must be sensitive to these differences.

In this study, we considered the cases and the controls as known groups and compared the values of the subscales in the two groups using the *t*-test for independent samples.

Reliability

Reliability can be defined as a measure of the internal consistency or homogeneity of the items: the more consistent the items, the greater is the reliability. The coefficient most frequently used to measure internal consistency is the well-known Cronbach's alpha (11). Cronbach's alpha values >0.80 indicate a reliable scale, although in the initial stages of a study, values >0.70 can be acceptable (12). Cronbach's alpha was calculated for each of the six OHIP subscales.

Results

A total of 124 patients (104 females, 20 males, mean age 35.1 years) constituted the study sample. According to the diagnostic criteria of the American Academy of Orofacial Pain (AAOP) (13), the TMD patients were divided into: one case of congenital and developmental disorders (0.8%), 35 of disc displacement with reduction (28.2%), 10 of disc displacement without reduction (8.1%), five of temporomandibular joint dislocation (4.0%), 12 of capsulitis/synovitis (9.7%), one of polyarthritides (0.8%), 21 of osteoarthritis primary (16.9%), 14 of myofascial pain (11.3%), six of myospasm (4.9%), 19 of local myalgia-unclassified (15.3%). We did not find any case of osteoarthritis secondary, ankylosis and fracture (condylar process) among the temporomandibular joint articular disorders and cases of myositis, myofibrotic contracture and neoplasia among the masticatory muscle disorders. Eighty-five (68.5%) subjects were affected by temporomandibular joint articular disorders, whereas 29 (31.5%) by masticatory muscle disorders. The 'control' group consisted of 61 subjects (45 females, 16 males, mean age 41.4 years).

Factorial structure

The number of eigenvalues >1, computed by exploratory factor analysis, totaled seven (Table 1). The percentage of cumulative variability explained by a model with six dimensions is 66%, whereas that explained by a model with seven dimensions is 70%. Therefore, given that the increment of

Table 1. Results of exploratory factor analysis

Dimension	Eigenvalue	% of variability	% of cumulative variability
1	9.47	32.65	32.65
2	3.76	12.98	45.63
3	1.84	6.35	51.97
4	1.60	5.50	57.48
5	1.32	4.54	62.02
6	1.15	3.96	65.98
7	1.07	3.69	69.66

explained variability attributable to the seventh dimension is low (3.68%) and as the seventh eigenvalue is very close to 1, we considered a six-factor model to explain the factorial structure of the data. The structure of the weights was, to a great extent, superimposable upon that identified by previous studies and reproduced the graphic layout of the original questionnaire.

Validity

As regards content validity, it is to be recalled that our questionnaire was obtained from an existing validated version, in English, and had, therefore, already undergone content analysis. Nevertheless, a group of expert dentists from the University of Pavia, Italy, re-analyzed the items and judged them to cover all the aspects of oral health-related quality of life. Following this content analysis, it was suggested that the item 'Been Self-conscious', be eliminated as most of the subjects did not understand its meaning.

In reference to criterion-related validation, Table 2 shows the Spearman correlation coefficients between the scores of the different subscales and the variable pain, which was the sum of the responses to questions regarding pain. The correlation coefficients revealed a moderate association for each subscale. All the coefficients were significantly different from 0 ($P < 0.05$).

Finally, the construction validity was investigated comparing the mean scores of each subscale recorded by the cases and controls. As shown in Table 3, the difference between the cases and controls was significant for all the subscales except 'social disability' and 'handicap'.

The constructs not showing statistical significance in the comparison between cases and controls were those related to situations that, evolving over time, indicated an advanced clinical situation. Therefore, taking the answers to the pain questions available for the cases, we subdivided the subjects into three groups: controls, initial cases, and advanced cases. Table 4 presents the mean scores on the subscales 'social inability' and 'handicap' in the advanced cases and controls. This time, the differences between the scores of the cases and the controls were statistically significant.

Table 2. Spearman correlation coefficients between the scores of the different subscales and the variable pain

	Pain (sum of D1–D10 scores)
Functional limitation	0.48
Psychologic discomfort	0.46
Physical disability	0.38
Psychologic disability	0.46
Social disability	0.32
Handicap	0.26

Table 3. Construction validity: comparison between mean scores of cases and controls

	Total mean score	Mean score for cases	Mean score for controls	*P*-value
Functional limitation	10.93	10.04	12.73	0.000
Psychological discomfort	11.88	11.32	13.09	0.004
Physical disability	28.52	27.71	30.32	0.003
Psychological disability	20.96	19.97	23.05	0.001
Social disability	21.33	21.11	21.76	0.289
Handicap	17.22	17	17.69	0.184
Number of subjects	185	124	61	

Table 4. Construction validity: comparison between controls and cases presenting an advanced clinical picture

	Mean score advanced cases	Mean score controls	*P*-value
Social disability	19.47	21.76	0.002
Handicap	15.70	17.69	0.003
Number of subjects	54	60	

Reliability

Table 5 shows the Cronbach's alpha values for each of the six subscales. The reliability of the psychological discomfort subscale is evaluated excluding the item 'Been Self-conscious'. The Cronbach's alpha value, always >0.70, shows quite good reliability for all six subscales. Considering that Cronbach's alpha values increase parallel with increases in the number of questionnaire items (14), obtaining alpha values >0.70 using three-/four-item scales is a very good result.

Table 5. Reliability: Cronbach's alpha values for each of the six subscales

	Cronbach's alpha
Functional limitation	0.72
Psychological discomfort	0.71
Physical disability	0.77
Psychological disability	0.86
Social disability	0.79
Handicap	0.79

Discussion and conclusions

We developed and validated an Italian version of the OHIP to measure the oral health impact profile in TMD patients. In the evaluation of quality of life, one of the WHO's 'health for everybody' fields, 'surveys' or questionnaires are considered a fundamental way of collecting data (15–17).

The questionnaire must be appropriate, comprehensible, nonequivocal, not subject to distortion, and capable of registering all possible answers, which must also be easy to codify. Furthermore, it must conform to relevant norms and standards, both judicial and ethical. At present, researchers wishing to undertake such a project have very little literature, especially in Italian, on which to draw from, and few examples on the basis of which to develop questionnaires.

As a result of the publication of health definition and measurement studies, the question of health-related quality of life began, in the mid-1980s, to be studied systematically. As this interest has been relatively recent, however, we had at our disposal very little up-to-date information on oral disease outcomes (18).

Early evaluations of oral health-related quality of life used general measurements, but these were unable to consider all the consequences of oral diseases. Evidence of the limits presented by these measures progressively led to the development of more specific and more suitable questionnaires. These are 'self-complete' questionnaires that investigate function, symptoms, and social and psychological impact (i.e. the three main quality-of-life aspects embraced by the multidimensional concept). These measurements are interesting for dentists because they evaluate how the patient perceives his oral health problems, why he opts for treatment, and how satisfied he is with the results. In this way, the practitioner can plan an effective and personalized course of action aimed at restoring, partly or totally, the patient's overall health.

We had to validate a questionnaire for use in Italy, and thus now have, at our disposal, a valid and reliable instrument that can be applied in future research. Analysis of the content validity of the questionnaire and of the various items led us to exclude the item 'Felt Self-conscious', as it was not understood by the most of the patients. We propose to eliminate this item from the final version of the questionnaire.

The six subscales showed an association with a pain measure that was taken as a reference variable in the study of criterion-related validity. The known groups' construct validity study revealed statistically significant differences between the mean score of the healthy subjects and that of the TMD patients in all the subscales except 'social disability' and 'handicap'. The content analysis of these two subscales showed they evaluate progressive aspects of the pathology that appear in chronic patients. This hypothesis was confirmed by the statistically significant differences that emerged between the mean scores of the healthy subjects and chronic patients.

In the reliability study, we found, for all six subscales, a Cronbach's alpha value >0.70, which indicates good internal consistency. Given that increases in the number of items in a questionnaire result in higher Cronbach's alpha values (14), obtaining alpha values >0.70 using three-/four-item scales must be considered a very good result.

The OHIP showed that the quality of life of subjects suffering from pain is definitely lower than that of pain-free subjects. So we can suppose that the patients' wellbeing decreases as a function of both pain duration and increases in pain intensity, frequency, and number of pain sites.

In fact, patients initially display functional limitations. These are followed by psychological discomfort, social disability and handicap, and finally chronic pain. This progression explains the different construct validity results displayed in Tables 3 and 4, which compare the controls with all the cases and with the advanced cases, respectively.

This validated instrument can be used in the diagnostic process to evaluate the quality-of-life impact of different treatments. Indeed, one of the newer objectives in the management (19) of TMD patients, in addition to reduction or elimination of pain and restoration of comfortable oral function, is to reduce their need for future treatments and thus to improve their quality of life.

Acknowledgements

The study was supported by Chiesi Farmaceutici SPA.

References

1. Locker D, Miller Y. Evaluation of subjective oral health status indicators. J Public Health Dent 1994;54:167–76.
2. Slade GD, Spencer AJ. Development and evaluation of the Oral Health Impact Profile. Community Dent Health 1994;11:3–11.

3. Locker D. Health outcomes of oral disorders. Int J Epidemiol 1995;24(Suppl 1):S85–9.
4. Slade GD, Spencer AJ. Social impact of oral conditions among older adults. Aust Dent J 1994;39:358–64.
5. Allison P, Locker D, Jokovic A, Slade G. A cross-cultural study of oral health values. J Dent Res 1999;78:643–9.
6. John MT, Patrick DL, Slade GD. The German version of the Oral Health Impact-translation and psychometric properties. Eur J Oral Sci 2002; 110:425–33.
7. Wong MC, Lo EC, McMillan AS. Validation of a Chinese version of the Oral Health Impact Profile (OHIP). Community Dent Oral Epidemiol 2002; 30:423–30.
8. Murray H, Locker D, Mock D, Tenenbaum HC. Pain and the quality of life in patients referred to a craniofacial pain unit. J Orofac Pain 1996;10:316–23.
9. 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;15:446–50.
10. Bartholomew D. Latent Variable Models and Factor Analysis. New York: Oxford University Press; 1987.
11. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16:297–334.
12. Nunnally JC. Psychometric Theory. New York: McGraw-Hill; 1978.
13. The American Academy of Orofacial Pain. Orofacial Pain. Guidelines for Assessments, Diagnosis, and Management. Chicago: Quintessence Publishing Co, Inc; 1996.
14. Hattie J. Methodology review: assessing unidimensionality of tests and items. Appl Psychol Measure 1985;9:139–64.
15. Locker D. Issues in measuring change in self-perceveid oral health status. Community Dent Oral Epidemiol 1998;26:41–7.
16. Reisine ST, Fertig J, Weber J, Leder S. Impact of dental conditions on patients' qualòity of life. Community Dent Oral Epidemiol 1989;17:7–10.
17. Reisine ST, Weber J. The effects of temporomandibular joint disorders on patients' quality of life. Community Dent Health 1989;6:257–70.
18. Locker D. Measuring oral health: a conceptual framework. Community Dent Health 1998;5:3–18.
19. Lund J, Lavigne G, Dubner R, Sessle B. Orofacial Pain. From Basic Science to Clinical Management. Chicago, IL: Quintessence Publishing Co, Inc; 2001.