# Assessing the Psychometric Characteristics of the Macedonian Version of the Oral Health Impact Profile Questionnaire (OHIP-MAC49)

Nikolina Kenig<sup>1</sup>, Julijana Nikolovska<sup>2</sup>

<sup>1</sup> PhD (Psychology), Associate Professor, Faculty of Philosophy.\* <sup>2</sup>PhD, DDS. Assistant Professor, Department for Prosthodontics, Faculty of Dental Medicine.\*

\* University of Ss. Cyril and Methodius in Skopje, Republic of Macedonia.

#### Abstract

Aim: The aim of this study was to adapt the Oral Health Impact Profile-49 (OHIP-49) for use by the Macedonian-speaking population and to assess its psychometric properties. Methods: After piloting a back-translated version of the OHIP-49 questionnaire in Macedonian, the resulting Macedonian version of the Oral Health Impact Profile questionnaire (OHIP-MAC49) was administered to 247 patients who had been recruited in four groups: Group 1 was composed of 163 randomly chosen blood donors representing the general population; Group 2 was a convenience sample of 20 patients who attended a clinic for relief of dental pain; Group 3 was a sample of 29 prosthodontic patients; and Group 4 was composed of 35 students. In order to ensure that all questions were answered, data were gathered from participants in the four groups by trained and experienced interviewers. The internal reliability of the OHIP-MAC49 scale and its constituent seven subscales was calculated for Groups 1, 3, and 4 by using Cronbach's alpha coefficient and average inter-item correlations. The test-retest stability of the instrument was estimated by calculating the intra-class correlation coefficients and the limits of agreement of the scores obtained from the participants in Groups 3 and 4 in a repeat interview three to four weeks after the first interview. Convergent validity was evaluated by comparing OHIP-MAC49 scores-both total (0-4) and subscores (2-4) with self-reported health for Group 3 patients—using the Spearman coefficient of correlation. For the purpose of evaluating group validity, the OHIP-MAC49 scores of Group 3 patients were compared by using Spearman's correlation coefficient. In addition, comparisons were made between patients with and without burning-mouth symptoms, temporomandibular pain, joint clicking, and oral habits (point-biserial correlation). The OHIP-MAC49 scores of subjects with and without dentures in Group 1 were compared by using point-biserial correlation. The responsiveness of the instrument was assessed by calculating the difference in OHIP-MAC49 means (total subscores) in Group 2 patients prior to and after treatment for relief of pain. Results: The internal consistency of each subscale and the whole scale estimated was excellent. Cronbach's alpha coefficients for whole scale ranged from 0.92 to 0.95. The intra-class correlation coefficients ranged from 0.83 to 0.99, suggesting that the instrument had satisfactory reliability in terms of time stability. The questionnaire had acceptable responsiveness, confirmed by a significant differences (P < 0.01) between the mean OHIP-MAC49 score at baseline and follow-up, both for the total sum (0-4) and the sum scores (2-4). Convergent validity, evaluated by comparing OHIP-MAC49 scores with self-reported oral health in Group 1 and Group 3 patients, was confirmed, because all correlation coefficients were significant (P < 0.01). The results from testing the anticipated differences on the basis of seven self-reported indicators of oral health in Group 3 patients, as well as the differences in OHIP-MAC49 scores between subjects in general population who wore and did not wear dentures, strongly suggest that the instrument has satisfactory group validity. Conclusion: The Macedonian version of the OHIP-49 demonstrated satisfactory validity, excellent reliability, and sufficient responsiveness and therefore can be used for assessing the impact of oral health on different aspects of quality of life in Macedonia.

Key words: Quality of Life, OHIP49, Macedonian Version, Psychometrics

## Introduction

Clinical indicators alone are insufficient to assess all oral health problems. The combination of clinical and subjective indicators provides a more comprehensive and multidimensional assessment of the patient's individual oral health status with resulting benefits for clinical decision making and oral health research [1-3]. To address the need to assess individuals' perception of their oral health status (subjective indicator), the Oral Health Impact Profile (OHIP) questionnaire was developed as a self-report instrument for assessing oral health-

Corresponding author: Julijana Nikolovska, Assistant Professor, Department for Prosthodontics, Faculty of Dental Medicine, University of Ss. Cyril and Methodius, Skopje, Republic of Macedonia; e-mail: julijananikolovska@yahoo.com

related discomfort and disability from several perspectives, including a patient's behavioural and social limitations. The theoretical background from which the indicators were derived is found in the conceptual framework proposed by the World Health Organization [4] and was adapted for oral health by Locker (1988) [5]. The original version of the OHIP consisted of 49 items, which are divided into seven subscales related to: (a) functional limitation, (b) physical pain, (c) psychological discomfort, (d) physical disability, (e) psychological disability, (f) social disability, and (g) handicap. Its reliability and sensitivity have been demonstrated in several studies [6,7] and its utility as an assessment tool has been shown in different cultures, after translation from English [8-13].

In the Balkan region, the translation and adaptation of the OHIP scale has been carried out in Croatia, Serbia, Romania, Turkey, and Greece [9-13], for either the long or the short version. It is now available in many languages, which makes it an excellent tool for conducting cross-cultural research in the realm of oral health-related quality of life (OHRQoL). However, to date it has not been adapted into a version for use in Macedonia. There was therefore a need to develop and test a Macedonian version (OHIP-MAC49).

## Aim

The aim of this study was to adapt the OHIP-49 for use by the Macedonian-speaking population and to assess its psychometric characteristics.

## Methods

#### Translation

The English version of the OHIP-49 was translated following the procedure that has been used in previous validation studies in other countries [8,14,15] and has been proposed in the guidelines for cross-cultural adaptation of self-reported instruments [16,17]. First, it was translated from English to Macedonian by two professional translators who were familiar with the terminology used in the questionnaire. They reported that it was not difficult to translate but they had to be cautious with several idiomatic expressions such as "painful aching", a "sense of taste", or feeling "self-conscious", simply because there were no exact expressions and few translations were possible. Because there were no significant differences in their translations, they were given to three Professors at the Department for Prosthodontics at the Faculty of Dental Medicine (University of Ss. Cyril and Methodius in Skopje) with excellent proficiency in English, to compare the original with the translated version. After they all independently agreed that the translation was accurate, the translated version was administered to a pilot sample of 22 subjects (patients) for the purpose of checking the general appropriateness of the wording and the instructions. The respondents had no difficulties in understanding the instructions or the questions, except for one item that was confusing, especially to those with less education. This question in the original version asked, "Have you been self-conscious because of your teeth, mouth or dentures?". Translated into Macedonian, it apparently had no semantic equivalence with the meaning in English, because half of the respondents needed further explanation in order to be able to understand it. This difficulty appeared to relate to the differences in the way the term "self-consciousness" is used in the two languages. In consultation with the translators, the question was rephrased and checked again with another ten patients. The modified version of this item worked well. The resulting translation was given to a bilingual (English and Macedonian) speaker who translated it back from Macedonian into English. There were no substantial differences from the original version to the one that was backtranslated and therefore the translation was considered to be adequate.

## **Participants**

The final translation of the questionnaire into Macedonian language was administered to a total number of 247 participants in four different groups. Group 1 (the general population) was represented by 163 randomly selected blood donors who were attending the Republic of Macedonia Institute for Transfusion in Skopje. The second group was composed of 20 patients who needed treatment for acute or chronic toothache and who attended the same clinic for relief of their oral pain. The third group was composed of 29 patients who were undergoing prosthodontic treatment at the University Dental Clinic Centre in Skopje. The fourth group was composed of 35 graduate dental students from the Faculty of Dental Medicine at University of Ss. Cyril and Methodius in Skopje. All participants used Macedonian as their first language. The description of the samples, along with the method of selection and the type of research that was conducted with each of them, is presented in Table 1.

Sample	N	Age in years	%	Sample	Type of assessment
		mean and (SD)	female	type	performed
Group 1:	163	34.7 (10.9)	45%	Random	Internal consistency
General population					and construct validity
Group 2: Patients	20	42.7 (14.0)	51%	Consecutive	Responsiveness
with demand for					(sensitivity to changes)
treatment					
Group 3:	29	59.9 (10.45)	59%	Convenience	Internal consistency:
Prosthodontic					Construct validity and
patients					test-retest
Group 4: Students	35	24.8 (1.6)	49%	Consecutive	Internal consistency
					and test-retest

Table 1. Overview of the sample (number, age and gender), sampling strategies and type of research used

To assess its test-retest reliability, the questionnaire was given twice to the same participants from the third and the fourth group, the second occasion being three to four weeks after the first. The Group 2 patients who had dental pain when they were first given the questionnaire were also tested twice for the purpose of assessing the responsiveness of the instrument. Both these procedures will be further explained in the sections "Reliability assessment" and "Responsiveness assessment", respectively.

## **Description of the OHIP-MAC49**

The original version of the instrument, the OHIP-MAC49 was composed of 49 core questions, divided into seven subscales. Respondents can choose the frequency for each of the conditions described in the questions on a five-point Likert-format scale, ranging from 0 = never, 1 = hardly, 2 = occasionally, 3 = fairly often, to 4 = very often. Taking into account that the questions represent a certain kind or aspect of impairment, a total score of zero  $(X_{min}=0)$  indicates total absence of any difficulty or problem, whereas higher scores suggest that the respondent experiences more oral health-related difficulties/problems. The highest possible score is  $X_{max}=196$  (49 x 4).

# Procedure for testing the OHIP-MAC49

The data were collected from March to May 2011. Ethical approval was obtained from the Ethical Committee of Faculty of Dental Medicine, University Ss. Cyril and Methodius in Skopje.

Each participant was given an explanation of the aim of the research, along with a guarantee of confidentiality. All responses were anonymous and made after the participants had given their informed consent. The average time taken to administer the OHIP-MAC49 was 30 minutes. Because all participants were interviewed, there were no missing data and the participation rate was 100%.

The respondents from Groups 1 and 3 were asked several additional questions that were used for assessing the construct validity of the instrument.

Prior to administering the questionnaire, respondents from Group 1 were also asked to determine their overall oral health on a five-point scale: excellent = 4, very good = 3, good = 2, fair = 1, and poor = 0, as well as stating whether or not they wore a denture.

## **Reliability assessment**

Two aspects of reliability were evaluated: the internal consistency (the homogeneity of the items) and test–retest stability (the stability of the scores over a reasonable period of time).

The internal consistency of the OHIP-MAC49 was tested by assessing Cronbach's alpha coefficients and average inter-item correlation coefficients for the whole scale and then for each of the seven subscales in Groups 1, 3 and 4 [18].

In order to assess the test-retest reliability, the questionnaire was administered twice to participants in Groups 3 and 4, with an interval of four weeks between the first and second administration. Prior to the second administration, all participants were asked whether they had received any dental treatment or experienced any dental problems since they first completed the OHIP-MAC49. They were given the retest only after they had confirmed that this was not the case.

Intra-class correlation coefficients (ICC) for the total scores and the seven subscores were calculated in order to evaluate the stability of scores between the first and second administrations of the questionnaire [19-20]. The ICCs were interpreted as follows: ICC <0.40 was considered as poor reliability, ICC between 0.40 and 0.75 was considered as fair to good reliability, and ICC higher than 0.75 was considered as an excellent reliability [21].

In addition, the difference in the total scores and subscores between the baseline and the followup was tested by the paired t-test. The limits of agreement around the mean difference (between the baseline and the follow-up means) were calculated within a confidence level of 95% [9,15,22].

## Validity assessment

As it is well documented in theory and practice, reliability is considered to be a necessary condition, but not a guarantee for the validity of instruments [23-25]. Therefore, the process of evaluating psychometric properties of OHIP-MAC49 should inevitably consider showing evidence that the instrument is valid; namely, that it measures what it intends to measure. Two aspects of construct validity were evaluated: *convergent validity* and *group validity*.

For the purpose of testing the convergent validity of the instrument, the association between the selfreported oral health score and the total score of the OHIP scale (OHIP 0-4) was examined. It was expected that the same associations would be present when answers that indicated total absence or very minor problems (0 = never and/or 1 = hardly ever) were excluded during the calculation of the total OHIP score (OHIP 2-4). If the hypothesis on the association between the OHIP score and the self-reported health is rejected, then the convergent validity of the scale should be considered as doubtful. Taking into account the level of measurement of the variables, the Spearman coefficient of correlation was deemed to be the most appropriate statistic for assessing the degree of association. The expected correlation between the self-reported oral health and the OHIP score was tested both for the general population (Group 1) and the prosthodontic patients (Group 3).

The *group* validity was assessed by testing the associations between several self-reported oral conditions; the total OHIP scores in Group 3 (prosthodontic patients) were tested. The four self-reported oral conditions were: (a) temporomandibular disorder (TMD) pain, (b) burning-mouth sensation, (c) joint clicking and (d) oral habits such as biting nails, lips or cheeks.

In accordance to what has been found in the other similar studies [8,9], it was expected that the presence of first three conditions would be associated with higher OHIP scores, whereas oral habits and joint clicking would not be significantly correlated with OHRQoL. Taking into account the format of these four additional questions, these hypotheses were tested by the means of point-biserial correlation.

It was also thought likely that three other oral conditions would be associated with the overall OHIP-MAC49 scores. These conditions were: (a) having a feeling that the denture does not fit well, (b) having pain and pressure caused by the denture, and (c) experiencing problems when eating during the last month. It was anticipated that the overall OHIP score is associated with higher reported degree on all of these conditions. Because the reported degrees are rank-order measurements, Spearman's coefficient of correlation was used to test these three hypotheses.

Finally, it was expected that there would be differences between participants in Group 1 (the general population) who wore dentures and those who do not. The hypothesis that there would be an association between denture wearing and the OHIP score was tested by assessing the point-biserial correlation between the two variables.

In order to strengthen the evidence regarding the conclusions for the hypotheses on the convergent and group validity, the correlations were checked with two summary OHIP sum scores: the total sum of items (0-4) and the sum of answers ranging from 2 to 4.

#### **Responsiveness assessment**

The responsiveness of the instrument can be defined as its ability to identify change in the variable that has been measured when it has occurred [26]. The responsiveness of this instrument was evaluated by testing its ability to detect the change in scores that was expected for the group of patients who requested treatment due to toothache. These subjects answered the OHIP-MAC49 questions twice: once immediately after the treatment; the second time, three to four weeks after the treatment. This was similar to the German and the Hungarian versions, where the period after the treatment was one month [27,28]. It was expected that their OHRQoL would improve after the treatment, taking into account that their painful tooth/teeth had been treated [9,22]. The difference between the first and the second assessments was calculated by the means of paired t-test. According to Cohen (1988), the effect size of 0.20 is considered small, 0.50 is considered medium, whereas 0.80 is regarded as being large [29].

#### Data analysis

The resulting data were entered into statistical software (Statistical Package for Social Sciences version 13 for Windows, SPSS Inc, Chicago, USA; Microsoft Office Excel 2007, Microsoft Corporation, Redmond, WA, USA). The statistical tests detailed previously in this methods section were applied.

#### Results

## **Reliability: internal consistency**

*Table 2* shows both Cronbach's coefficients and the average inter-item correlations as statistical indicators of the internal consistency of the instrument. Each of the subscales has high internal homogene-

ity, especially for Group 1 (general population) and Group 3 (prosthodontic patients) because almost all á-values were higher than 0,70. When administered to Group 4 (students), three subscales (social disability, handicap, and functional limitation) had lower internal consistency. However, Cronbach's alpha coefficients for the whole scale showed very high consistency in measuring the construct, because they ranged from 0.92 to 0.95.

The average inter-item correlations also suggested that the internal consistency of the scale was satisfactory. It is worth noting that the majority of subscales had the most desirable inter-item correlation that ranged between 0.40 and 0.50; namely, the level suggesting that there were no redundant items that caused high reliability coefficients (*Table 2*).

## **Test-retest reliability**

As can be seen from *Table 3*, the interclass correlation coefficients were high in both groups that were

 Table 2. Internal consistency of the OHIP-MAC49 and its seven subscales (Cronbach's alpha coefficients and average inter-item correlations)

	Group 1: General population (N=163)		Gro Pros	up 3: thodontic	Group 4: Students		
			patien	nts 1 (N=29)	(N=35)		
OHIP subscales	Cronbach Inter-item		Cronbach	Inter-item	Cronbach	Inter-item	
(n of items)	alpha	correlation	alpha	correlation	alpha	correlation	
Functional limitation (9)	0.91	0.56	0.64	0.15	0.59	0.16	
Physical pain (9)	0.83	0.40	0.69	0.31	0.88	0.50	
Psychological discomfort (5)	0.88	0.63	0.66	0.20	0.77	0.50	
Physical disability (9)	0.91	0.55	0.79	0.35	0.77	0.24	
Psychological disability (6)	0.86	0.52	0.79	0.40	0.84	0.47	
Social disability (5)	0.81	0.47	0.85	0.58	0.50	0.20	
Handicap (6)	0.84 0.47		0.74	0.38	0.59	0.25	
OHIP-MAC49	0.95 0.35		0.92	0.22	0.94	0.35	

 Table 3. Test-retest reliability of OHIP-MAC49 and its subscales measured by intraclass correlation coefficients (ICC)

	Group 3: Prosthodontic patients			Group 4: Students			
		(N=29)		(N=35)			
Scale (n of items)	ICC	Means of	Limits of	ICC	Means of	Limits of	
		differences	agreement		differences	agreement	
Functional limitation (9)	0.92	0.96*	0.08 to 1.84	0.99	-0.03	-0.16 to1.03	
Physical pain (9)	0.94	3.45**	2.45 to 4.45	0.99	-0.17	-0.40 to 0.57	
Psychological discomfort (5)	0.94	1.24*	0.57 to 1.91	0.97	0.86	0.48 to 1.23	
Physical disability (9)	0.86	-1.21	-2.31 to -0.11	0.98	-0.14	-0.40 to 0.11	
Psychological disability (6)	0.66	0.24	0.08 to 0.40	0.98	-0.11	-0.36 to 0.13	
Social disability (5)	0.97	0.41	-0.01 to 0.83	0.99	0.03	-0.29 to 0.09	
Handicap (6)	0.98	-0.10	-0.48 to 0.27	0.97	-0.06	-0.20 to 0.09	
OHIP-MAC49	0.98	-5.17**	-7.68 to 7.68	0.99	0.34	-0.34 to 1.03	

\*\**P*<0.01, \* *P*<0.05

retested: Group 3 (prosthodontic patients) and Group 4 (students). Apart from the subscale "psychological disability" for the group of prosthodontic patients, ICCs were all higher than 0.80, which is regarded as being excellent reliability in terms of time stability [30].

## **Construct validity**

The results of the statistical tests of the proposed hypothesis regarding the convergent and group validity are shown in *Tables 4*, 5 and 6. The convergent validity of the scale was verified by the findings that the Spearman correlations between the self-reported health (on a scale ranging from 0 = unsatisfactory to 4 = excellent) and the OHIP-MAC49 (0-4), as well as OHIP-MAC49 (2-4) sum scores were statistically significant at the 0.01 level (*Table 4*). The expected general assumption that the subjective assessment of the oral health is connected with the overall OHIP score was confirmed.

Additionally, the point-biserial correlations between the OHIP-MAC49 scores (0-4) and the 2-4 sum scores and denture wearing were statistically significant (P<0.01) and confirmed the expectation that subjects who wore a denture would have higher OHIP scores in comparison with those who did not (*Table 5*).

Furthermore, the group validity was tested by examining the associations among the OHIP-MAC49 scores and several indicators of oral health in Group 3 (prosthodontic patients) (*Table 6*). Except in the case of reported pain and pressure caused by the denture, where the correlation was not significant only when the sum scores were calculated by excluding the answers indicating absence of problems (2-4), all observed correla-

 

 Table 4. Convergent validity assessment: correlation between OHIP-MAC49 scores and self-reported oral health

	N (%)	OHIP	OHIP	Correlation and
		0-4 (M)	2-4 (M)	significance
Group 1: General population (N	=163)			
Self-reported oral health				-0. 528** (OHIP 0-4)
				-0. 560**(OHIP 2-4)
excellent	76	10.28	16.59	
very good	41	24.73	10.53	
fair	31	35.42	19.42	
poor	10	50.70	40.35	
unsatisfactory	5	59.40	30.86	
<b>Group 3: Prosthodontic patients</b>	(N=29)	·		
Self-reported oral health				-0.804** (OHIP 0-4)
				-0.684** (OHIP 0-4)
poor	5	111.00	79.20	
fair	6	59.00	42.67	
good	14	42.43	37.43	
very good	2	35.50	29.00	
excellent	2	13.50	9.00	

\*\*P<0.01

 Table 5. Groups validity assessment: correlation between denture wearing and OHIP-MAC49 scores in general population

Group 1: General population	N (%)	OHIP	OHIP	Correlation and	
(N=163)		0-4 (M)	2-4 (M)	significance	
Denture				0.225** (OHIP 0-4)	
				0.239** (OHIP 2-4)	
Yes	6	70.50	61.83		
No	157	22.30	14.86		

\*\*P<0.01

Group 3: Prosthodontic patients (N=29)	Ν	OHIP 0-4 (M)	OHIP 2-4 (M)	Correlation and significance	
Feeling that the denture				0.436* (OHIP 0-4)	
does not fit well				0.475 **(OHIP 2-4)	
never	7	33.29	24.57		
hardly	4	61.00	50.75		
occasionally	8	62.13	46.00		
fairly often	4	57.25	48.00		
very often	6	66.33	54.33		
Pain and pressure caused				0.384* (OHIP 0-4)	
by the denture	0	20.90	22.67	0. 351 (OHIP 2-4)	
	9	52.50	33.07		
hardly	0	55.50	42.17		
	5	60.00	50.00		
Tairly often	4	62.25	41.50		
very often	5	/4.40	57.80		
Experiencing problems				0.515** (OHIP 0-4)	
when eating		22.20	26.00	0.527** (OHIP 2-4)	
never	5	32.20	26.00		
hardly	9	48.11	37.33		
occasionally	4	41.50	40.75		
fairly often	9	82.22	60.33		
very often	2	50.50	44.50		
Temporomandibular disorder pain				0.364* (OHIP 0-4) 0.365* (OHIP 2-4)	
Yes	5	92.20	69.60		
No	24	47.50	38.04		
Burning-mouth sensation				0.427** (OHIP 0-4) 0.475** (OHIP 2-4)	
Yes	8	82.63	59.25	()	
No	21	44.76	37.48		
Oral habits				-0.007 (OHIP 2-4) 0.047 (OHIP 0-2)	
Yes	7	51.86	38.00		
No	22	56.27	45.23		
Joint clicking				0.107 (OHIP 2-4) 0.164 (OHIP 0-2)	
Yes	2	59.00	42.67		
No	27	54.77	43.58		

 

 Table 6. Group validity assessment: correlation between self-reported oral health indicators and OHIP-MAC49 scores in prosthodontic patients

\*\**P*<0.01, \*P<0.05

tions between the proposed self-reported oral health indicators and self-reported oral conditions with OHRQoL followed the predicted association. The majority of predicted correlations were significant at the 0.01 level. Moreover, in cases where associations were not anticipated (oral habits and joint clicking), no significant correlations were observed ( $r_{pbis}$ = -0.007 for OHIP 2-4,  $r_{pbis}$ = 0.047

for OHIP 2-4,  $r_{pbis}$ = -0.107 for OHIP 2-4, and  $r_{pbis}$ = 0.164 for OHIP 2-4). All these findings suggest that the OHIP-MAC49 had satisfactory construct validity.

#### Responsiveness

In Group 2 (patients with treatment demand from acute and chronic pain), the average OHIP-MAC49

		<b>OHIP (0-4)</b>	<b>OHIP (2-4)</b>
Group 2: Patients with	M (baseline)-M (follow-up)	55.00-38.45**	45.20-24.75**
treatment demand	95% confidence interval	6.21-26.89	10.43-30.46
for pain (N=20)	Effect size (Cohen)	0.57	0.63

Table 7. Responsiveness of the OHIP-MAC49 scale

\*\*P<0.01

score (0-4) decreased from 55.0 to 38.5 after the treatment. As can be seen in *Table 7*, there was a significant difference between the baseline and the follow-up means of the total OHIP (0-4) sum score and for the total OHIP (2-4) sum score (P<0.01). In both cases, Cohen s d indicated medium effect of means differences. Because it was expected that the follow-up mean would decrease significantly after the treatment, these results can be regarded as a confirmation of the responsiveness of the instrument.

## Discussion

The sampling strategy used in this study was similar to the one employed in the assessment of the psychometric properties of the Croatian, German, and Hungarian OHIP versions [9,27,28]. The decision to select the particular four groups was made not only for practical reasons, but also because it was comparable to the aforementioned three studies. The size of the samples is also very similar to those in these studies, especially if it is taken into account that the overall population size in Macedonia is much smaller. However, because there were only 29 prosthodontic patients, it would be desirable to repeat this part of the study that evaluated internal consistency with a larger sample.

Although the full OHIP-MAC49 questionnaire has many questions and its administration in the form of interview was time consuming, the vast majority of participants were willing (and some of them even enthusiastic) to answer it. The participants compliance strengthens the credibility of the results and also suggests that the questions were understandable and relevant, and that the adaptation was well done. Its administration, scoring and interpretation are reasonably easy and the whole procedure does not require any special training that goes beyond the basic principles.

The evaluation of the internal consistency of the total OHIP-MAC49 and its seven subscales was made by using the method employed elsewhere for this purpose. The Cronbach s alpha coefficients for the whole scale (0.92-0.95) suggest that the scale had excellent internal consistency and they were not very different from those found in other studies that have assessed the reliability of the 49-item scale [8,9,27,31]. Such high internal consistency cannot be ascribed merely to the length of the instrument, although it might seem so if it is considered that the majority of short versions have somewhat lower internal consistency coefficients [15,32,33], because the majority of the average inter-item correlations are in the recommended range of 0.40 0.50. In this study, three average inter-item correlation were higher than 0.50 in the general population and only one in the group of prosthodontic patients, indicating possible redundancy of items within the particular subscales. Future research might resolve this doubt through the use of factor analysis, as has already been undertaken [34].

In previous studies [9,14], some of the scales had lower coefficients of internal consistency than the limit that is proposed as being desirable. In the current study, this was the case with three different subscales in the groups of prosthodontic patients and students. This was not surprising, taking into account that both groups were smaller in size than the group from the general population and even more importantly, quite homogeneous. It might be expected that for the group of prosthodontics patients, the acquired Cronbach s alpha coefficients would have been larger if the sample were more heterogeneous [35].

The test retest reliability of the total OHIP-MAC49 score and the seven subscales was satisfactory for both tested study samples and the findings, which was not surprising, taking into account that it is a common finding for the 49-item version of the OHIP as well as for the short ones [8].

The Macedonian version of OHIP can be considered as being internally responsive. There was a significant difference between the baseline and the follow-up means of the total OHIP (0-4) sum score and for the total OHIP (2-4) sum score (P<0.01). Cohen s d indicated medium effect of means differences, lower than in the Croatian study [9] where it was assessed in the same way, and somewhat higher than the Slovenian shorter version [22]. However, the moderate effect size is not unusual in cases where the samples of patients with treatment demand are heterogeneous in terms of the aetiology of pain. When the majority of patients have acute dental pain, unlike those in the current study, the effect size can be expected to be larger [22].

The correlation coefficients between the OHIP-MAC49 scores and the self-reported oral health in general population strongly suggest that the instrument has acceptable convergent validity [23,26], already confirmed in virtually all other studies that employed comparison between scores on self-reported health and the total OHIP scores [8,9].

The associations between the OHIP-MAC49 scores and the several proposed self-reported indicators of oral health in prosthodontic patients strongly suggest that the scale has well-supported construct validity. The expected differences between the groups were confirmed statistically and in cases when they were not expected, they were not found. The findings regarding the same hypotheses (experiencing temporomandibular pain, burning-mouth sensation, oral habits, and joint clicking) are replicated [9]. Only one hypothesis was not confirmed (the association between the OHIP-MAC49 score and experiencing pain and pressure caused by a denture) but only when the scores were calculated by excluding the answers that indicated absence of symptoms. One possible explanation for this finding might be that the participants tended to give answers indicating absence of problems caused directly by the denture.

This finding suggests yet another limitation of the study that needs to be addressed by future research. The indicators used to evaluate the validity of the instrument predominantly rely on selfreported measurements [22] that might be compromised by response bias [23]. Future research should use more objective (clinical) variables, which

# References

1. Slade GD, Spencer AJ. Development and evaluation of the Oral Health Impact Profile. *Community Dental Health*. 1994; **11**: 3-11.

2. Allen PF. Assessment of oral health related quality of life. *Health and Quality of Life Outcomes.* 2003; 1: 40.

3. Locker D. Issues in measuring change in self-perceived oral health status. *Community Dentistry and Oral Epidemiology*. 1998, **26**: 41-47.

4. World Health Organization (WHO). International Classification of Impairments, Disabilities and Handicaps. Geneva: WHO; 1980.

5. Locker D. Measuring oral health: a conceptual framework. *Community Dental Health.* 1988; **5**: 3-18. should bring about more reach and more objective information on the validity of the instrument.

Taking into account the evaluated psychometric properties, it could be inferred that the OHIP-MAC49 version meets most of the desired attributes for self-report instruments [36]. Therefore, it can be used as a valid and reliable instrument for assessing the subjective evaluation of the oral health-related quality of life in Macedonia.

## Conclusion

The Macedonian version of the OHIP-49 demonstrated satisfactory validity, reliability, and responsiveness and therefore can be used for assessing the impact of oral health on different aspects of quality of life in Macedonia.

## Acknowledgement

We thank all the participants for their time and appreciate the suggestions and support of Dr. Nikola Petricevic, University of Zagreb, School of Dentistry, Croatia. The authors thank the editorial team of this journal for their help and advice.

# Contributions of each author

- NK contributed by performing the statistical analysis and interpreting the findings, and writing the Abstract, Methods, Results (including Tables), and Discussion sections.
- JN contributed by initiating the study, performing the field work (interviews), performing data entry, and writing the Introduction and References sections.
- Both authors read and approved the final manuscript.

# Statement of conflict of interest

As far as the authors are aware, there is no conflict of interests.

6. Allison P, Locker D, Jokovic A, Slade G. A cross-cultural study of oral health values. *Journal of Dental Research*. 1999; **78**: 643-649.

7. Montero-Martín J, Bravo-Pérez M, Albaladejo-Martínez A, Hernández-Martín LA, Rosel-Gallardo EM. Validation the Oral Health Impact Profile (OHIP-14sp) for adults in Spain. *Medicina Oral Patologia y Cirugia Bucal.* 2009; **14**: E44-50.

8. Meulen MJ, John MT, Naeije M, Lobbezoo F. The Dutch version of the Oral Health Impact Profile (OHIP-NL): Translation, reliability and construct validity. *BMC Oral Health.* 2008; **8**: 11.

9. Petricevic N, Celebic A, Papic M, Rener-Sitar K. The Croatian version of the Oral Health Impact Profile questionnaire. Collegium Antropologicum. 2009; 3: 315-321.

10. Stančić I, Tihaček Šojić Lj, Jelenković A. [Adaptation of Oral Health Impact Profile (OHIP-14) index for measuring impact of oral health on quality of life in elderly to Serbian language]. *Vojnosanitetski Pregled*. 2009; **66**: 511-515. [Article in Serbian]

11. Murariu A, Hanganu C. Oral health and quality of life among 45- to 64-year-old patients attending a clinic in Iasi, Romania. Oral Health and Dental Management in the Black Sea Countries. 2009; **8**(2): 7-11.

12. Caglayan F, Altun O, Miloglu O, Kaya MD, Yilmaz AB. Correlation between oral health-related quality of life (OHQoL) and oral disorders in a Turkish patient population. *Medicina Oral, Patologia y Cirugia Bucal.* 2009; **14**(11): e573-578.

13. Roumani T, Oulis CJ, Papagiannopoulou V, Yfantopoulos J. Validation of a Greek version of the oral health impact profile (OHIP-14) in adolescents. *European Archive of Paediatric Dentistry*. 2010; **11**: 247-252.

14. Lopez R, Baelum V. Spanish version of the Oral Health Impact Profile (OHIP-Sp). *BMC Oral Health.* 2006; **6**: 11.

15. Larsson P, List T, Lundström I, Marcusson A, Ohrbach R. Reliability and validity of a Swedish version of the Oral Health Impact Profile (OHIP-S). *Acta Odontologica Scandinavica.* 2004; **62**: 147-152.

16. Domino M, Domino ML. *Psychological Testing: An Introduction.* 2nd ed. Cambridge: Cambridge University Press; 2006.

17. Beaton DE, Bombardier C, Guillemin F, Marcos Bosi Ferraz M. Guidelines for the process of cross-cultural adaptation of self-report measures. *SPINE*. 2000; **25**: 3186-3191.

18. Bland JM, Altman DG. Cronbach's alpha. British Medical Journal. 1997; **314**: 572.

19. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*. 1999; **8**: 135-160.

20. Shrout PE, Fleiss J. Intra-class correlations: uses in assessing rater reliability. *Psychological Bulletin.* 1979; **86**: 420-428.

21. Fleiss JL. The Design and Analysis of Clinical Experiments. New York, NY: Wiley; 1986.

22. Rener-Sitar K, Petričević N, Čelebić A, Marion L. Psychometric properties of Croatian and Slovenian short form of Oral Health Impact Profile questionnaires. *Croatian Medical Journal.* 2008; **49**: 536–544. 23. Urbina S. *Essentials of Psychological Testing;* Hoboken, NJ: John Wiley & Sons; 2004.

24. Kaplan RM, Saccuzzo, DP. Psychological Testing: Principles, Applications, and Issues. 2nd ed. Pacific Grove, CA: Brooks/Cole; 1989.

25. Sawilowsky SS. Reliability: Rejoinder to Thompson and Vacha-Haase. In: Thompson B, editor. *Score Reliability: Contemporary Thinking on Reliability Issues.* Thousand Oaks, CA: Sage; 2003. p. 149-154.

26. Roach KE. Measurement of health outcomes: reliability, validity and responsiveness. *Journal of Prosthetists and Orthotists.* 2006; **18**(1S) :8-12.

27. John MT, Patrick DL, Slade GD. The German version of the Oral Health Impact Profile: translation and psychometric properties. *European Journal of Oral Sciences*. 2002; **110**: 425-433.

28. Szentpetery A, Szabo G, Marada G, Szanto I, John MT. The Hungarian version of the Oral Health Impact Profile. *European Journal of Oral Sciences*. 2006; **114**: 197-203.

29. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

30. John MT, Miglioretti DL, LeResche L, Koepsell TD, Hujoel P, Micheelis W. German short forms of the Oral Health Impact Profile. *Community Dentistry and Oral Epidemiology*. 2006; **34**: 277-288.

31. Oliveira BD, Nadanovsky P. Psychometric properties of the Brazilian version of the Oral Health Impact Profile: short form. *Community Dentistry and Oral Epidemiology*. 2005; **33**: 307-314.

32. Navabi N, Nakhaee N, Mirzadeh A. Validation of a Persian version of the Oral Health Impact Profile (OHIP-14). *Iranian Journal of Public Health.* 2010; **39**: 135-139.

33. Fernandes MJ, Ruta DA, Ogden GR, Pitts NB, Ogstone SA. Assessing oral health-related quality of life in general dental practice in Scotland: validation of the OHIP-14. *Community Dentistry and Oral Epidemiology*. 2006; **34**: 53-63.

34. John MT, Hujoel P, Miglioretti DL, LeResche L, Koepsell TD, Micheelis W. Dimensions of oral-health-related quality of life. *Journal of Dental Research*. 2004; **83**: 956-960.

35. Traub RE. *Reliability for the Social Sciences Theory and Applications*. Vol 3. Thousand Oaks, CA: Sage; 1994.

36. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research.* 2002; **11**: 193-205.