

RESEARCH

Open Access



The use of Subject Matter Experts in Validating an Oral Health-Related Quality of Life measure in Korean

Jaesung Seo¹, Michael MacEntee² and Mario Brondani^{3*}

Abstract

Objectives: This paper aimed to employ subject matter experts (SMEs) to assess the extent to which the Korean version of the short-form of the OHIP (OHIP-14 K) is culturally valid and equivalent in Korean.

Methods: We approached 17 bilingual Korean SMEs from which 10 independently rated the clarity, relevance, and cultural equivalence of the OHIP-14 K. SME's varied between 10 and 41 years of clinical experience and were mostly males (# 7). We used Item-level Content Validity Index (I-CVI) to gauge the proportion of SMEs who considered the content of OHIP items (e.g., instruction, response format, etc.) to be culturally valid. We also performed additional analysis to determine the level of agreement between the SMEs.

Results: The experts rated most of the items to be clear (S-CVI = 0.93) while having difficulties in assigning relevance of the questions to the expected domains (S-CVI = 0.42). Moreover, considerable disagreement existed among the experts in regard to the relevance (Kfree = 0.19 to 1.00) and the cultural equivalence indexes (ADM = 0.36 to 0.96). The content of the OHIP-14 K for the most part clearly reproduced the language of the original OHIP-14. However, experts disagreed on the relevance and conceptual equivalence of the OHIP-14 K for a Korean population.

Conclusions: Patient-oriented outcome measures such as the OHIP can be used across cultures once there are indeed assessing the same domains and constructs of interest. The CVI technique seems to be an alternative tool for evaluating content validity and equivalency of an OHQoL measure. A more refined, culturally relevant version of OHIP-14 K was proposed although there is no available data yet to support a better score validity, reliability and responsiveness of this proposed version.

Introduction

Oral health-related quality of life (OHQoL) represents a psychological *construct* defined as self-reports pertaining to the functional, psychological and social impacts of oral problems on quality of life [1]. Many OHQoL measures are available worldwide for exploring the self-perceived status of oral health via surveys and for comparing before and after treatment outcomes via clinical trials, for example [2–6]. The Oral Health Impact Profile (OHIP) has been the most used self-reported measure of OHQoL based on the International Classification of Impairment,

Disability and Handicap as interpreted to oral health by David Locker in 1988 [7–9]. It consists of 49 questions representing seven domains (Table 1) assessed by a 5-point Likert response scale (“very often”; “fairly often”; “occasionally”; “hardly ever”; or “never”, with an optional “don’t know”). The OHIP-14 is a shortened version of the OHIP-49 reduced through regression and item-impact analysis to 14 questions [10].

Adapting a QoL measure like the OHIP from English to another language offers the possibility for cross-cultural comparisons [11–13] once assumed that its conceptual foundation in the ICIDH and Locker’s model is accepted at face value, and that the concepts and domains addressed are readily transferable between cultures. However, cultural environment strongly influence personal identity and how people consider, interpret and cope with

* Correspondence: Brondani@dentistry.ubc.ca

³Department of Oral Health Sciences, Division of Preventive & Community Dentistry, and Prosthodontics and Dental Geriatrics, University of British Columbia, JBM 122/2199 Wesbrook Mall, Vancouver, BC V6T 1Z3, Canada
Full list of author information is available at the end of the article

Table 1 The theoretical domains and functional items of the original short form of the Oral Health Impact Profile (OHIP-14)

Theoretical Domains	OHIP-14 Item
Functional Limitation	Q1. Have you had trouble pronouncing any words because of problems with your teeth, mouth or dentures?
	Q2. Have you felt that your sense of taste has worsened because of problems with your teeth, mouth or dentures?
Pain & Discomfort	Q3. Have you had painful aching in your mouth?
	Q4. Have you found it uncomfortable to eat any foods because of problems with your teeth, mouth or dentures?
Psychological Discomfort	Q5. Have you been self-conscious because of your teeth, mouth or dentures?
	Q6. Have you felt tense because of problems with your teeth, mouth or dentures?
Physical Disability	Q7. Has your diet been unsatisfactory because of problems with your teeth, mouth or dentures?
	Q8. Have you had to interrupt meals because of problems with your teeth, mouth or dentures?
Psychological Disability	Q9. Have you found it difficult to relax because of problems with your teeth, mouth or dentures?
	Q10. Have you been a bit embarrassed because of problems with your teeth, mouth or dentures?
Social Disability	Q11. Have you been a bit irritable with other people because of problems with your teeth, mouth or dentures?
	Q12. Have you had difficulty doing your usual jobs because of problems with your teeth, mouth or dentures?
Handicap	Q13. Have you felt that life in general was less satisfying because of problems with your teeth, mouth or dentures?
	Q14. Have you been totally unable to function because of problems with your teeth, mouth or dentures?

chronic diseases and disorders [5, 6]. The OHIP has been translated into more than 30 languages via various methodologies, but not without challenges to achieve validity and equivalency to the original Australian-English version [6]. Despite reservations about the validity of the OHIP's conceptual foundations [5,6] and domains' structure [14], Bae et al. (2007) adapted the OHIP-14 to Korean (OHIP-14 K) using bilingual translators without fully validating its concepts; yet with assumptions that it is a valid and equivalent translation [15]. The objective for this study was to assess the content validity and cultural equivalence of the Korean version of the OHIP-14 K with the assistance of bilingual subject matter experts (SMEs).

Methods

Selection of Subject Matter Experts - Ethics, consent and permissions

Ethical approval was obtained from the University of British Columbia Behavioral Research Ethics Board # H10-01023. The selection of expert participants was based on the 2009 College of Dental Surgeons of British Columbia directory by purposefully searching for practicing dentists with Korean first and family names. The participants included those who were practicing in Vancouver at least three times a week for a minimum of 10 years, understood and used both the Korean and English, and were willing to participate in this study. Although the number of recruited SMEs varies [16, 17], Beck and Gable (1986) suggested a minimum of 10 participants to yield acceptably consistent responses and to avoid chance agreement (Table 2). [18] We have identified 20 eligible participants; 17 were successfully contacted as the other 3 could not be reached at their listed addresses. From the 17 who were contacted, 10

volunteered to participate while the others refused to take part in the study because of their busy schedules or lack of interest..

Consent to publish

Consent to publish was obtained from the participants as they signed off on the following statement on the written consent form: *"Your signature indicates that you consent to participate in this study. You are willing to have your interview audio-taped and give permission for the principal investigator to use the information you are providing anonymously as part of a publication focused on the same issue."*

Content validation (CV) began with the construction and administration of a questionnaire to gather quantitative and qualitative information on the clarity, cultural equivalency and relevance of the OHIP [14,18], and was pilot-tested for clarity by two local practicing Korean dentists outside the group of 10 who participated in the study. They both recommended that the WHO's definitions of impairment, disability, and handicap (1980) be appended as background material to consult.

We then designed a questionnaire with three comprehensive subscales for SMEs to judge: 1) the content validity of OHIP-14 K in terms of the technical quality of the items, instructions, and response formats; 2) the relevance of the content to the ICDH theoretical domains of Locker's model; and 3) the cultural equivalency of the translation to the intent of the original OHIP-14. On the recommendation of Lynn [16], the CV questionnaire acquired a 4-point ordinal Likert scale as a response format without neutral ground.

In the clarity index, the SMEs were asked to rate the clarity of the instructions, items, and response format

Table 2 Background information of the subject matter experts

SME	Gender	Place(s) of Graduation	Clinical Experience (yrs)	Education Background	Location
1	M	University of British Columbia; University of Temple; University of California, San Francisco	15	DDS, PhD, AEGD ¹	Burnaby, BC
2	M	Seoul National University; University of British Columbia	26	DDS/MD	Burnaby, BC
3	M	Undisclosed	30	DMD	Coquitlam, BC
4	M	Yonsei University; University of Manitoba; University of British Columbia	19	DMD, MSc, PhD	Burnaby, BC
5	F	Southwestern University	14	DMD	North Vancouver, BC
6	M	University of London; Seoul National University; Korea University	41	BDS, DMD, MSc, PhD	Vancouver, BC/ Seoul, Korea
7	F	University of British Columbia	10	BDS, DDS	Burnaby, BC
8	M	University of Manitoba	15	DDS	North Vancouver, BC/ Lancaster, CA
9	F	University of Pennsylvania; Columbia University	12	DDS, MA	Coquitlam, BC
10	M	Korea University; University of British Columbia	26	DDS (SNU), DMD (UBC), MSD, PhD, PPD	Burnaby, BC

¹Advanced Education in General Dentistry

using the following Likert scale: 0 = not at all clear, 1 = somewhat clear, 2 = mostly clear, and 3 = very clear [19]. They were instructed to identify comprehension problems with the OHIP-14 K due to, for example: vague wording, ambiguous language, double-barrelled questions, and so on.

In the cultural equivalence index, the SMEs were asked to evaluate the semantic, colloquial, experiential and conceptual equivalence, rather than linguistic or literal equivalence of the translation using the scale: 0 = not at all equivalent, 1 = somewhat equivalent, 2 = mostly equivalent, and 3 = equivalent. They were encouraged to also comment generally on the translation with suggestions for deletions, additions and modifications.

In the relevance index, the degree to which the OHIP-14 K items appropriately sampled the theoretical domains of OHIP (e.g., functional limitation, social disability, etc.) was assessed. The SMEs were asked to identify for each item the most appropriate theoretical domain of Locker’s model.

All information, including the various types of cultural equivalence, was written and verbally conveyed to the SMEs when meeting with the first author (JS). After this introductory briefing, the SMEs signed the informed consent form and received a CV questionnaire with the instructions and the OHIP-14 in both Korean and English. They answered the CVI questionnaires individually, independently, and at their own convenience. The SMEs were again visited within 30 days where the questionnaires were collected.

Data analysis

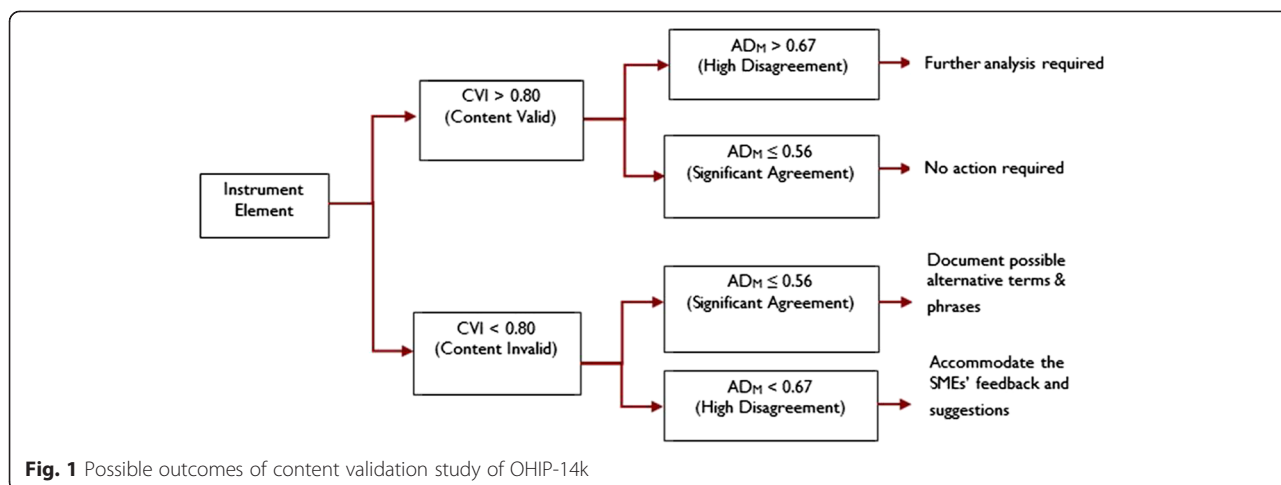
The Likert responses were analyzed using SPSS® (IBM Corp, Version 19.0. Armonk, NY: IBM Corp). The clarity

of the OHIP-14 K was rated for relevance and equivalence both for the entire scale (S-CVI) and individually for the instructions, response format, event frequency, and items (I-CVI). The I-CVIs were calculated as the proportion of SMEs who endorsed the validity of each scale (i.e., ratings of 3 or 4 on the 4-point Likert scale) as suggested by Beck and Gable [18].

$$CVI = \frac{\text{Number of responses as “3 or “4”}}{\text{Total number of responses}}$$

The AD_M measures multi-rater disagreement in the scale’s units and uncovers hidden disagreement in dichotomous data. AD_M for the clarity and cultural equivalence indices was calculated as the sum of the differences between individual ratings and the mean in absolute values divided by the total number of ratings. A lower AD_M value indicated stronger agreement between SME-ratings because the AD_M is dispersed around the mean, while disagreement might lead to revisions on the instrument (Fig. 1).

We used non-parametric Kappa statistics for the analysis of data generated by the relevance index in which the most representative theoretical domain was selected as suggested by Slocumb and Cole [20]. Kappa values below the cut-off 0.4 were considered to be poor agreement [20] and prompted further examination of the SMEs’ comments. The Scale-level CVI was expressed as the percentage of items whose I-CVI values were equal to or greater than the minimally acceptable CVI of 0.78; it provided information on the proportion of elements requiring revisions until the S-



CVIs was equal to or greater than 0.80 to confirm the validity of the scale [21].

Results

Clarity of OHIP-14 K Elements

With the exception of the response format, all 16 elements (including items and OHIP instructions) had I-CVI values greater than 0.80, indicating adequate levels of clarity (Table 3). The deviation from the mean index (AD_M) across all elements was below the critical value of 0.56, suggesting that homogeneity in SME ratings was unlikely to have been achieved by chance. For the most part, the OHIP-14 K was judged to be clear ($S-CVI = 0.93$, $AD_M \leq 0.56$). The only element whose CVI value felt below the acceptable level was the response format ($I-CVI = 0.7$), which was considered to be vague. SMEs recommended changing the Korean words “maewoo” (very) to “maewoo jaju” (very often), and “guhee” (hardly) to “guhee junhyu” (hardly ever). In addition, three other SMEs commented that the Korean translation of the OHIP-14 asks about symptoms that overlap significantly with systemic illnesses such as depression, so it could potentially have diverse interpretations. Consequently, they recommended that the specific oral health context be expressed in every question by including the phrase “because of problems of your mouth, teeth or dentures” as in the original English version.

Cultural Equivalence between OHIP-14 and OHIP-14 K

The SMEs were instructed to evaluate the cross-cultural equivalence between the OHIP-14 in English (OHIP-14E) and OHIP-14 K. Seven elements – which included the response format and questions 7, 8, 9, 10, 12, and 13 – were deemed content valid ($CVI > .80$), all with acceptable agreement ($AD_M \leq 0.56$). On the contrary, items including the instructions, the frequency scale, and questions 5 and 6 felt below the minimally

acceptable CVI value with statistically significant agreement ($I-CVI < 0.80$, $AD_M \leq 0.56$). While neither significant agreement nor disagreement was observed for items 2 and 4, a high level of disagreement was noted for items 1, 3, 11, and 14 with regard to cultural equivalency. For example, the SMEs were divided over the cultural equivalency of the Korean translation of question 1: *trouble pronouncing words* ($AD_M = 0.96$). Four SMEs suggested that *having trouble pronouncing words* had a different meaning than the Korean translation *discomfort from not being able to pronounce well* as offered in the current OHIP-14 K. The suggested revisions for question 1 included “*baleumeul mothaesuh himdeushinjuk*” (having difficulties to pronounce [any words]). The experts also disagreed over the equivalency of question 14 ($AD_M = 0.72$), “*totally unable to function,*” which was translated into “*jungshinjeok, shinchejeok, sahoejeokeuro junhyuh jemokeul halsu upsutdun jeok*” (totally unable to do one’s share psychologically, physically, and socially).

Relevance of OHIP-14 K Domains

Table 4 shows the distribution of SME’s responses when asked “Which one of the domains best represents each OHIP question?” and the descriptive statistics including each item’s CVI value, K_{free} , and levels of agreement.

Examination of each question’s relevance to the expected theoretical domain revealed that the endorsement rates for questions 1, 2, 3, 6, 11, and 14 were equal to or above the acceptable CVI value of 0.80 ($S-CVI_{relevance} = 0.42$, $K_{free} > 0.4$), confirming that the questions corresponded with the hypothesized domains by the Locker’s model. On the other hand, eight out of 14 questions (# 4, 5, 7, 8, 9, 10, 12, and 13) felt below the acceptable CVI value, with 5 showing poor or fair agreement ($K_{free} < 0.4$). Closer examination of content-invalid items revealed that question 5, “*self-conscious*”

Table 3 Item-level content Validity Index (I-CVI) and Average Deviation from the Mean Index (AD_M) of OHIP-14 K instruction, response format, event frequency, and items in the Clarity and Equivalence Indices

Original Scale Element (<i>back-translation of OHIP14-K</i>)	Clarity Index ¹		Equivalence Index ²	
	I-CVI	AD _M	I-CVI	AD _M
Instruction	0.9	0.38	0.3	0.55
Response Format	0.7	0.34	0.8	0.63
Event Frequency ^a	1.0	0.40	0.6	0.54
1. Trouble pronouncing any words (<i>Discomfort from not being able to pronounce well</i>)	1.0	0.40	0.6	0.96
2. Sense of taste has worsened (<i>Sense of taste was worse than before</i>)	1.0	0.38	0.7	0.64
3. Painful aching in your mouth (<i>Pain in the tongue, sublingual, cheeks, palate, etc.</i>)	1.0	0.40	0.4	0.70
4. Uncomfortable to eat any foods (<i>Uncomfortable to have meal due to painful or uneasy problems of the mouth</i>)	1.0	0.39	0.7	0.60
5. Self-conscious (<i>Reluctant to meet others because of shame</i>)	0.9	0.38	0.4	0.56
6. Felt tense (<i>Paid attention to</i>)	0.9	0.38	0.6	0.54
7. Unsatisfactory diet (<i>Dissatisfied meals</i>)	1.0	0.40	0.9	0.48
8. Meals interrupted (<i>Interrupted during meals</i>)	1.0	0.40	0.8	0.36
9. Difficult to relax (<i>Difficulties resting comfortably</i>)	1.0	0.39	0.8	0.36
10. A bit embarrassed (<i>Embarrassed or perplexed</i>)	0.9	0.39	0.9	0.36
11. A bit irritable with other people (<i>Get angry easily at others</i>)	0.9	0.37	0.5	0.70
12. Difficulty doing your usual jobs (<i>Difficult to do normal jobs</i>)	0.9	0.37	0.9	0.56
13. Life in general was less satisfying (<i>Life less satisfying than before</i>)	0.9	0.37	0.8	0.54
14. Totally unable to function (<i>Psychologically, physically and socially cannot at all do one's share</i>)	0.8	0.36	0.6	0.72

^aTwo missing values from two questionnaires (SME 7 and 8) were imputed with the item mean value for that scale element.

¹Clarity Index S-CVI = 0.93 ²Equivalence Index S-CVI = 0.50

(psychological discomfort); question 9, “*difficult to relax*” (psychological disability); and question 13, “*life in general was less satisfying*” (handicap) were also representing social disability ($K_{free} = 0.4$), physical disability ($K_{free} = 0.6$), and psychological disability ($K_{free} = 0.4$), respectively. Poor agreement was noted for questions 7, 10, and 12 ($K_{free} < 0.2$) while questions 4 and 8 showed fair agreement ($0.2 < K_{free} < 0.4$). Overall, 7 SMEs indicated that some questions could be interpreted outside of the oral health context.). For example, question 12, “*difficulties in*

doing one's usual jobs,” may unintentionally elicit non-dental-related experiences if left ‘as is’. The seemingly transferrable construct of OHQoL can be understood differently in English and in non-western cultures such as Korean due to differences in priorities, health perceptions and potential impact of a disorder. Another SME indicated that OHIP-14 K does not adequately capture the aesthetic concerns that patients might have about their mouths. In their overall evaluation of OHIP-14 K, 6 SMEs recognized the need for better

Table 4 The SMEs' classification of items into the seven OHIP domains, Item-level Content Validity Index (I-CVI), Free-Marginal Kappa (K_{free}), and levels of agreement on the Relevance Index

Item	Numbers of SMEs who endorsed each domain as the most representative for that item							Descriptive Statistics		
	Functional Limitation	Physical Pain	Psychological Discomfort	Physical Disability	Psychological Disability	Social Disability	Handicap	I-CVI	K_{free}	Agreement
Q1	8	-	-	-	1	1	-	0.80	0.56	Moderate
Q2	9	-	-	-	1	-	-	0.90	0.77	Substantial
Q3	-	10	-	-	-	-	-	1.00	1.00	Perfect
Q4	6	-	-	3	1	-	-	0.00	0.30	Fair
Q5	-	-	5	-	-	5	-	0.50	0.35	Fair
Q6	-	-	10	-	-	-	-	1.00	1.00	Perfect
Q7	-	-	4	4	2	-	-	0.40	0.17	Poor
Q8	2	-	2	6	-	-	-	0.60	0.27	Fair
Q9	-	-	2	8	-	-	-	0.00	0.59	Moderate
Q10	-	-	5	-	3	2	-	0.30	0.19	Poor
Q11	-	-	1	-	1	8	-	0.80	0.56	Moderate
Q12	2	-	-	6	-	-	2	0.00	0.20	Poor
Q13	-	-	4	-	6	-	-	0.00	0.38	Fair
Q14	-	-	-	-	1	-	9	0.90	0.77	Substantial

Note: Shaded cells indicate the theoretical domains predetermined by Slade and Spencer (1994).

accuracy of the OHIP-14 K items to ensure cross-cultural equivalence.

Discussion

Our study employed SMEs to culturally assess the content and equivalence of an OHRQoL instrument to Korean. As advised by Sisco and Broder (2011), OHQoL has multiple applications in dental research and services especially when we move from a bench research to a more person-centered approaches to measure treatment needs and efficacy of care [22]. In turn, the availability of cross-culturally valid and reliable OHQoL measures is beneficial for needs assessment, oral health care planning, and

service evaluation in Korean as well as in other languages. Despite the widespread use of OHIP in English and in more than 30 different languages, content validity and equivalency of its translated versions – including the OHIP-14 K – have not been fully addressed [6]. This was compounded by the fact that current cultural adaptation and validation strategies using the suggested forward and backward translations supervised by a committee are not resistant to biases; ramifications to inferences made on cultural differences in OHQoL are expected. Typical validation efforts for the OHIP use criterion-related approaches that are vulnerable to the cross-cultural biases and misunderstandings while paying little attention to the content of

the scale or theoretical foundations. Consequently, the validity of the OHIP translated to other languages and applied to other cultures needs a critical discussion [5] since none of the existing translations of the OHIP, including the Korean, seems to challenge Locker's ICIDH concept of disturbances to oral health-related quality of life [6].

In contrast to the traditional committee method of establishing equivalency [23–25], our study employed SMEs who investigated the theoretical foundations of the OHIP-14 K and recommended changes. The scale-level CVIs obtained in our study demonstrated to be an alternative method for content validation and indicated that wording of the OHIP-14 K is clear ($S-CVI_{\text{clarity}} = 0.93$), but it may not be cross-culturally equivalent to its English counterpart ($S-CVI_{\text{equivalence}} = 0.50$). The positive results on the clarity index were expected considering that OHIP-14 K has already undergone rigorous testing with monolingual Korean adults and five Korean dentists [15] despite its cultural equivalency never being fully established. Our results also indicated a limited degree of relevance for the entire scale ($S-CVI_{\text{relevance}} = 0.42$). Cross-culturally valid scales must demonstrate evidence of item relevance and mutual exclusiveness of their theoretical dimensions to achieve proper representation of the construct [26]. For this reason, each of the seven domains should be represented by at least two content-valid items to avoid taking chances with translation quality. Only half of 14 OHIP-14 K questions accurately loaded onto the expected seven domains as proposed by Slade and Spenser, while John and colleagues employed exploratory factor analysis to determine that the OHIP lends itself to four, not seven domains [14]. The low level of relevance of the set of assigned domains can imply a departure from the original conceptual model and an inaccurate representation of the construct since the subscales are not measuring what they are supposed to measure. This finding also raises questions about the use of subscale scores as a valid and reliable indicator of OHQoL domains while little attention has been placed to discuss responsiveness to changes of scores from the translated OHIP in the clinical status of respondents [5, 14].

The scale elements met the minimally acceptable disagreement level ($AD_M \leq 0.56$ or $K_{free} > 0.40$) but were below the acceptable 0.8 value of CVI and needed to be revised or eliminated according to the experts' suggestions. However, instead of eliminating items from an already short scale, revisions of the OHIP-14 K instructions, response format and frequency as well as questions 5, 6, and 9 were suggested (Table 5). Standardizing the recall periods is necessary to minimize construct-irrelevant variance for making valid cross-

cultural comparisons. In addition, the inclusion of the phrase “because of problems with your teeth, mouth or dentures”, omitted in the Korean version [15], would help respondents to focus on oral health-related events when answering the questions.

In the literature, various types of biases have been reported in other language versions of the OHIP. When question 3, “have you had any painful spots or areas in your mouth?” was translated into Brazilian Portuguese, it used the word “pontos” for spot. However, the word also has a second meaning – suture stitches – which caused confusion among the respondents [27]. A similar translation problem was reported by Kenig and Nikolovska (2012) in the Macedonian version of the OHIP item 5, “self-consciousness” [28]. In Macedonian, the literal translation of the term had a different meaning than intended whereas in the Korean version, the low degrees of translation equivalence ($I-CVI = 0.4$) and relevance ($I-CVI = 0.5$) suggest that the translation may have been too liberal. In the case of the Macedonian version, the authors decided to eliminate the item because most respondents did not understand its meaning [29].

There are a number of possible explanations for the reported translation problems. As concluded by Guillemín et al. [12, 25], there is no “standardized approach to the cross-cultural adaptation of HRQOL instruments” which probably corroborates the fact that there is no detailed explanation of how conceptual equivalence has been indeed explored within the OHIP [14]. Another source of asymmetry could have been the ambiguities in the English version itself. As discussed previously, the *handicap* domain included questions that were artificially added to the original OHIP and not directly developed from the interviews with the lay Australian respondents [8]. In the case of our study, none of the SMEs judged question 13, “life less satisfying,” to represent the handicap domain. Moreover, disagreement was noted for question 14, “totally unable to function,” which was translated to signify “a state of being incapacitated psychologically, physically and socially”. Not only can such a triple-barrelled question be confusing for respondents, but it also gives a very vague impression of “handicap” for which no equivalent word exists in Korean. Cultural equivalence aside, these two questions were loaded into psychosocial impact, one of the four domains identified by John et al. to better structure the OHIP and to be ‘similar across cultures and populations’ [14]. Likewise, our study found that seemingly equivalent items did not always guarantee their relevance to the expected theoretical domains. For example, although question 9, “have you had difficulties relaxing?” was judged to be equivalent to its English counterpart ($I-CVI_{\text{equivalence}} = 0.8$), it did not load onto the

Table 5 Suggested revised version of OHIP-14 K

지난 1년 동안 아래의 경험을 얼마나 자주 겪으셨습니까?	매우 자주	자주	가끔	거의 전혀	전혀
1. 치아, 구강이나 틀니 때문에 발음이 잘 안되어 불편했던 적이 있으십니까?					
2. 치아, 구강이나 틀니 때문에 맛을 느끼는 감각이 예전보다 나빠졌다고 느끼신 적이 있습니까?					
3. 혀 밑, 뺨 입천정 등이 아픈 적이 있습니까?					
4. 치아, 구강이나 틀니 때문에 아프거나 거북스러운 입안의 문제 때문에 음식 먹기가 불편한 적이 있습니까?					
5. 치아, 구강이나 틀니 때문에 신경이 많이 쓰인 적이 있습니까?					
6. 치아, 구강이나 틀니 때문에 긴장이 되신 적이 있습니까?					
7. 치아, 구강이나 틀니 때문에 식생활이 불만스러운 적이 있습니까?					
8. 치아, 구강이나 틀니 때문에 식사를 도중에 중단하신 적이 있습니까?					
9. 치아, 구강이나 틀니 때문에 긴장을 풀지 못하신 적이 있습니까?					
10. 치아, 구강이나 틀니 때문에 난처하거나 당황스러웠던 적이 있습니까?					
11. 치아, 구강이나 틀니 때문에 다른 사람들에게 화를 잘 내게 되신 적이 있습니까?					
12. 치아, 구강이나 틀니 때문에 평소 하시던 일을 하기가 어려웠던 적이 있습니까?					
13. 치아, 구강이나 틀니 때문에 살아가는 것이 예전에 비해서 덜 만족스럽다고 느끼신 적이 있습니까?					
14. 치아, 구강이나 틀니 때문에 정신적 신체적 사회적으로 전혀 계 문을 할 수 없었던 적이 있습니까?					

expected *psychological disability* domain ($I-CVI_{relevance} = 0.0$). A similar translation issue was reported by Room and colleagues (1996), who noted that Korean translations of psychological affective states were easily mistaken for physical states because Korean words regarding feelings do not effectively differentiate between physical sensations and emotions [30].

Finally, there could have been conceptual differences across cultures in interpreting the semantic equivalence as advised by Herdman [31] who highlighted that translated instruments implicitly and explicitly assume that notions of oral health-related quality of life are

similar across cultures, when they may not be. Hence, while revisions of the questions are critical to the validity of cross-cultural comparisons of OHQoL, conceptual equivalence remains the most challenging part of the translational process and it has been either addressed superficially or omitted altogether from many of the OHIP translations [6].

Implications of the Content Validity Findings

Our study found limited evidence of content validity for OHIP-14 K in terms of relevance and cultural equivalency with the English version. In line with

previous research [32, 33], the SMEs' suggestions for improving the scale's content validity underscored the importance of its cultural appropriateness and faithfulness to the source version and theoretical mode. Our study suggests that OHIP-14 K should have the same relationship with the construct of interest both within and across cultural groups.

Another implication of our study includes the impact of the detected biases on cross-cultural comparisons of OHQoL. The same degree of construct might elicit different responses on a Likert scale and consequently biased interpretation of domain and total scores [34]. For example, a physical disability domain could be measuring social disability, or could be loaded into a structured factor to better characterise OHQoL across cultures [14]. Hence, the fragile relationship between items and their theoretical domains carries significant implications as OHIP scores are calculated for each domain as well as for the entire scale [5]. However, Brondani and MacEntee raised a more fundamental problem with the use of a summative score due to the questionable discreteness and stability of the theoretical domains of Locker's model. While John et al. (2014) suggested four factors to better structure the OHIP domains [14], Bakers (2007) questioned whether or not the OHIP domains were actually distinguishable [35] and if so, how they would readily relate to one another as per Locker's conceptual framework.

Based on the findings presented here, a refined version of OHIP-14 K was yielded to enhance both semantic and conceptual equivalence (Table 5). Although more work is needed to evaluate its psychometric properties in a target Korean sampling group, it emerged based on the assumption that the existing OHIP Korean translation holds validity. Although there is no question that the OHIP is a psychometric instrument tested and widely used, there are other techniques to provide insight about the OHIP's ability to measure invariance across populations.

Limitations of our study include the use of a convenient sample of dentists who suggested changes in the OHIP as an exercise of content validity and cultural equivalency. Multi-disciplinary SMEs from other disciplines could have provided a more diverse range of knowledge and experience as conducted by others [17, 36]. The CVI technique used should be supplemented with other validation testing specially within the target population. Likewise, the proposed OHIP version on Table 5 remains critical to be consulted by lay Koreans who could examine the relevance and utility of the OHIP-14 K revisions suggested in this study. In turn, no data exists to support a better scoring system for validity, reliability and responsiveness in the proposed modified version of the OHIP to the target population.

Conclusion

Like the OHIP, patient-oriented outcome measures can only enhance our appreciation for the relationships between oral and general health across cultures once they are indeed assessing the same domains and constructs of interest within a content valid and culturally equivalent measure. Our study showed that:

- The CVI technique is an alternative tool for evaluating content validity and equivalency of an OHQoL measure and documenting the content validation process and quantification of CVIs and disagreement indices.
- The expert suggestions and the CVI ratings could be used also to improve the content validity and equivalency of the OHIP-14 K, as well as to refine inferences made from cross-cultural measurements of OHQoL.
- The OHIP-14 K demonstrated limited evidence of content validity and cultural equivalency, and potential cross-cultural biases have been identified in its method, items, and construct representation.
- Future studies should be done to establish content validity and cultural equivalency of other language versions of OHIP-14 (or other OHQoL scales) and further explore the utility of CVIs and disagreement indices as there is a need for a continuous evaluation of the scale for the intended target populations.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JS proposed and carried out the study, interviewed the SMS's, and analysed the data. MB drafted the manuscript and formatted the text. MM edited the earlier drafts of this manuscript and gave editorial suggestions. All authors read and approved the final manuscript as submitted.

Acknowledgments

This manuscript was originated as part of the first author (JS) requirements for the completion of a Masters of Dental Sciences, which was successfully concluded in July, 2012. We would like to thank the SMEs who donated their time to participate in our study. A special 'thank you' goes to Drs. Ross Bryant, Peter Graf and Sabrina Wong for their guidance and encouragement.

Author details

¹Schulich School of Medicine and Dentistry Western University, Ontario, Canada. ²Department of Oral Health, Sciences. Division of Prosthodontics and Dental Geriatrics, University of British Columbia, Vancouver, Canada. ³Department of Oral Health Sciences. Division of Preventive & Community Dentistry, and Prosthodontics and Dental Geriatrics, University of British Columbia, JBM 122/2199 Wesbrook Mall, Vancouver, BC V6T 1Z3, Canada.

Received: 10 March 2015 Accepted: 26 August 2015

Published online: 04 September 2015

References

1. Gift HC, Redford M. Oral health and quality of life. *Clin Geriatric Med.* 1992;8:673–83.
2. Slade GD. Oral Health Impact Profile. In: Slade GD, editor. *Measuring Oral Health and Quality of Life.* Chapel Hill, NC: Dental Ecology; 1997. p. 94–104.

3. Jokovic A, Locker D, Stephens M, Kenny D, Tompson B, Guyatt G. Validity and Reliability of a Questionnaire for Measuring Child Oral-health-related Quality of Life. *J Dent Res*. 2002;81(7):459–63.
4. Hebling E, Pereira AC. Oral health-related quality of life: A critical appraisal of assessment tools used in elderly people. *Gerodontology*. 2007;24(3):151–61.
5. Brondani MA, MacEntee MI. The concept of validity in sociodemographic indicators and oral health-related quality-of-life measures. *Comm Dent Oral Epidemiol*. 2007;34:472–8.
6. MacEntee MI, Brondani MA Cross-cultural Equivalency in Translations of the OHIP. *Comm Dent Oral Epidemiol*, 2015, in press.
7. Locker D. Measuring oral health: a conceptual framework. *Comm Dental Health*. 1998;5:3–18.
8. Slade GD, Spencer AJ. Development and evaluation of the oral health impact profile. *Comm Dent Health*. 1994;11(1):3–11.
9. John MT, Hujoel P, Miglioretti DL, LeResche L, Koepsell TD, Micheelis W. Dimensions of Oral-health-related Quality of Life. *J Dent Res*. 2004;83(12):956–60.
10. 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(5):446–50.
11. Hambleton RK, Patsula L. Adapting tests for use in multiple languages and cultures. *Social Indicators Res*. 1998;45(1):151–71.
12. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health related quality of life measures: literature review and proposed guidelines. *J Clinical Epidemiol*. 1993;46:1417–32.
13. Brondani M, He S. Translating Oral Health-Related Quality of Life Measures: Are There Alternative Methodologies? Translating quality of life measures. *Social Indicators Res*. 2010;106(2):1–15.
14. John MT, Reissmann DR, Feuerstahler L, Waller N, Baba K, Larsson P, et al. Exploratory factor analysis of the oral health impact profile. *J Oral Rehab*. 2014;41:635–43.
15. Bae KH, Kim HD, Jung SH, Park DY, Kim JB, Paik DI, et al. Validation of the Korean version of the oral health impact profile among the Korean elderly. *Comm Dent Oral Epidemiol*. 2007;35(1):73–9.
16. Lynn MR. Determination and quantification of content validity. *Nursing Res*. 1986;35:382–5.
17. Domingues GB, Gallani MC, Gobatto CA, Miura CT, Rodrigues RC, Myers J. Cultural adaptation of an instrument to assess physical fitness in cardiac patients. *Rev Saude Publica*. 2011;45(2):276–85.
18. Beck CT, Gable RK. Ensuring content validity: An illustration of the process. *J Nurs Measur*. 2001;9(2):201–15.
19. Grant JS, Davis LL. Selection and use of content experts for instrument development. *Res Nursing Health*. 1997;20(3):269–74.
20. Ferketich S. Aspects of item analysis. *Res Nurs Health*. 1991;14:165–8.
21. Slocumb EM, Cole FL. A practical approach to content validation. *Applied Nurs Res*. 1991;4(4):192–5.
22. Sischo L, Broder HL. Oral Health-related Quality of Life: What, Why, How, and Future Implication. *J Dent Res*. 2011;90(11):1264–70.
23. Devriendt E, Van den Heede K, Coussement J, Dejaeger E, Surmont K, Heylen D, et al. Content validity and internal consistency of the Dutch translation of the Safety Attitudes Questionnaire: An observational study. *Inter J Nursing Studies*. 2012;49(3):327–37.
24. Harkness J, van de Vijver FJR, Johnson T. Questionnaire Design in Comparative Research. In: Harkness J, van de Vijver FJR, Mohler PP, editors. *Cross-Cultural Survey Methods*. New York: Wiley; 2003. p. 19–34.
25. Guillemin F. Cross-cultural Adaptation and Validation of Health Status Measures. *Scandinavian J Rheumatology*. 1995;24:61–3.
26. Smit J, Van den Berg CE, Bekker LG, Stein DJ, Seedat S. Translation and cross-cultural adaptation of a mental health battery in an African setting. *Afr Health Sci*. 2006;6:215–22.
27. Pires CPAB, Ferraz MB, Abreu MHNG. Translation into Brazilian Portuguese, cultural adaptation and validation of the oral health impact profile (OHIP-49). *Brazilian Oral Res*. 2006;10(3):263–8.
28. Kenig N, Nikolovska J. Assessing the psychometric characteristics of the Macedonian version of the Oral Health Impact Profile questionnaire (OHIP-MAC49). *Oral Health Dent Manag*. 2011;11(1):29–38.
29. Segu M, Collesano V, Lobbia S, Rezzani C. Cross-cultural validation of a short form of the Oral Health Impact Profile for temporomandibular disorders. *Comm Dent Oral Epidemiol*. 2005;33(2):125–30.
30. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Qual Life Res*. 1998;7:323–35.
31. Room R, Janca A, Bennett LA, Schmidt L, Sartorius N. WHO cross cultural applicability research on diagnosis and assessment of substance use disorders: An overview of methods and selected results. *Social Addiction*. 1996;91:199–220.
32. Alegria M, Vila D, Woo M, Canino G, Takeuchi MV, Febo V, et al. Cultural relevance and equivalence in the NLAAS instrument: Integrating etic and emic in the development of cross-cultural measures for a psychiatric epidemiology and services study of Latinos. *Intern J Methods Psych Res*. 2004;13(4):270–88.
33. Cenoz J, Todeva E. The well and the bucket: the emic and etic perspectives combined. In: Todeva E, Cenoz J, editors. *The Multiple Realities of Multilingualism*. Berlin: Mouton de Gruyter; 2009. p. 265–92.
34. Anderson RT, McFarlane M, Naughton MJ, Shumaker SA. Conceptual issues and considerations in cross-cultural validation of generic health-related quality of life instruments. In: Spilker B, editor. *Quality of life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996. p. 605–12.
35. Bakers SR. Testing a Conceptual Model of Oral Health: a Structural Equation Modeling Approach. *J Dental Res*. 2007;86:708–12.
36. King KM, Khan N, Leblanc P, Quan H. Examining and establishing translational and conceptual equivalence of survey questionnaires for a multi-ethnic, multi-language study. *J Advanced Nurs*. 2011;67(10):2267–74.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

