Validation of the English basal and squamous cell carcinoma quality of life (BaSQoL) questionnaire

Wesley Y. Yu*
Rick Waalboer-Spuij*
Rebecca Bremer
Brian Lu
Christine Aroyan
Lauren Crow
Roy Grekin
Isaac Neuhaus
Siegrid Yu
Sarah T. Arron
Loes M. Hollestein

*contributed equally to this paper.
Accepted Dermatol Surg.

ABSTRACT

Background: Keratinocyte carcinomas (KC) impact patient quality of life (QoL). There is a need for validated QoL instruments specific to KC. The Basal and Squamous Cell Carcinoma Quality of Life (BaSQoL) questionnaire was developed to comprehensively measure issues of importance to patients with KC.

Objective: To validate and characterize the BaSQoL questionnaire for QoL measurement following diagnosis and treatment of KC.

Methods: This was a prospective, observational study. Patients with basal or squamous cell carcinoma were asked to fill out BaSQoL, Skin Cancer Index (SCI), and Hospital Anxiety and Depression Scale (HADS) questionnaires. Descriptive statistics and classical test theory were used to assess validity.

Results: 187 subjects enrolled in this study: 122 with BCC and 65 with SCC. 171 subjects (91.4%) completed questionnaires at all three time points; 16 patients (8.6%) were lost to follow up. Overall performance using classical test theory was good, with good internal consistency (Cronbach's s 0.63-0.80). BaSQoL subscales were strongly correlated with subscales of the SCI, demonstrating convergent validity, and weakly correlated with HADS, showing divergent validity.

Conclusion: The English language version of BaSQoL has good face, content, and construct validity. This study validates BaSQoL for use in English-speaking patients with BCC and SCC.

INTRODUCTION

The keratinocyte carcinomas (KC), basal and squamous cell carcinoma (BCC, SCC), severely impact patient quality of life (QoL). [1] Patients may suffer from pain, bleeding, social embarrassment, and anxiety surrounding the diagnosis of cancer. Objective and accurate tools to measure these experiences and document the impact of KC are necessary. [2] General QoL tools, such as the Hospital Anxiety and Depression Scale, are not targeted towards skin disease. Within the field of dermatology, the patient experience of skin cancer is different from that of inflammatory skin disease. Skin-targeted quality of life questionnaires such as the Dermatology Life Quality Index (DLQI) are not specific enough to demonstrate significant quality of life impairment in patients with skin cancer, and show little to no improvement in quality of life after treatment. [3,4] Skindex, another general dermatologic questionnaire, is not specific enough to measure the impact of KC on quality of life. [5] More targeted questionnaires such as the Skin Cancer Index (SCI) and the Skin Cancer Quality of Life Impact Tool (SCQOLIT) have advanced the field but have limitations. [6,7] For example, neither captures one of the most reported issues in skin cancer patients—the necessary behavioral changes regarding sun exposure. [8] In addition, the SCI does not capture anxiety about the treatment itself, other than scarring. [6] Finally, neither tool addresses the full spectrum of dermatologic issues such as the burden of frequent skin checks, triggering worries about other skin diseases, and the behavior change necessary to prevent future skin cancers. [4, 6-9]

The Basal and Squamous Cell Carcinoma Quality of Life (BaSQoL) questionnaire was developed through a rigorous multi-phase process to comprehensively measure problems specific to patients with keratinocyte carcinomas, such as fear of scars and coping mechanisms, worries about subsequent skin tumors, and the burden of sun protective behavior. [8,10] As described previously, topics in the questionnaire were generated through exhaustive patient focus groups led by independent psychologists and semi-structured interviews with healthcare providers. These items were then reviewed by an interdisciplinary expert panel (including dermatologists, dermatologic surgeons, plastic surgeons, general practitioners, ophthalmologists, and head-neck ear nose and throat (ENT) surgeons) then presented to patients for feedback. Patients and physicians were asked to rate items for inclusion in the study. Finally, the questionnaire was field tested in 1,173 patients selected from the Netherlands Cancer Registry and the questionnaire was reduced using exploratory factor analysis and item response theory. In this study, we validate the English translation of the BaSQoL, compare its performance to the SCI and HADS, and demonstrate the utility of BaSQoL in measuring quality of life in patients with skin cancer before and after surgical treatment.

METHODS

The study was conducted from July 1, 2017 to June 30, 2018. Study data were collected and managed using Research Electronic Data Capture (REDCap) tools hosted at University of California, San Francisco (UCSF). [11] A consecutive sample of patients with the diagnosis of SCC or BCC and scheduled for treatment in the UCSF Dermatologic Surgery center were approached for voluntary participation. Demographic and clinical information, including age, sex, contact information, tumor type, tumor size, tumor location, history of skin cancer, and treatment plan were collected. Participants were not compensated. The UCSF Institutional Review Board approved the study and all participants provided written informed consent.

The original BaSQoL questionnaire consisting of 16 questions written in Dutch was translated to English using forward-backward translation. [12,13] The BaSQoL consists of 5 subscales (behavior, diagnosis & treatment, worries, appearance and other people) scored from 0 to 3, with higher scores indicating poorer QoL. Participants were asked to fill out this questionnaire one week before their treatment (T0), four weeks after treatment (T1), and five weeks after treatment (T2). Additionally, participants were asked to complete the Skin Cancer Index (SCI) and the Hospital Anxiety and Depression Scale (HADS) one week before treatment (T0). The SCI consists of three subscales (emotional, social, appearance) with standardized score ranges from 0 to 100, with higher scores indicating higher QoL. We hypothesized high correlations between comparable subscales of the SCI and the BaSQoL (convergent validity). The HADS consists of two subscales (anxiety and depression) with scores ranging from 0 to 21, with higher scores indicating poorer QoL. The HADS was included to demonstrate divergent validity.

We aimed to include at least 100 patients with BCC and 50 with SCC in order to include sufficiently large samples of both types of KC, recognizing that BCC is the more common tumor. A formal sample size calculation was not performed. Classical test theory is a framework for evaluating the reliability of items in a questionnaire. Eight performance features (item difficulty, response distribution, item-test correlation, item-rest correlation, discriminant validity, item complexity, internal consistency, stepwise regression) were tested for each item on the BaSQoL (Table S1). Descriptive statistics were used to test item difficulty (missing responses) and response distribution. Spearman's correlation coefficients were calculated for item-test and item-rest correlation, and to test item discriminant validity. Internal consistency was tested via Cronbach's α coefficients. A forward stepwise regression was performed for each subscale in order to check the percentage of variance explained by the items in a subscale. The multitrait-multimethod correlation matrix was used to assess convergent and discriminant validity. We accounted for multiple

hypothesis testing where appropriate by correcting p-values using the false discovery rate (FDR) calculated by the Benjamini-Hochberg procedure. [14] In order to test the stability of BaSQoL responses over time, T1 and T2 responses were compared using a two-way mixed effect model to calculate the intraclass correlation coefficients. Analyses were performed in IBM SPSS Statistics for Windows, Version 21.0 (Armonk, New York: IBM Corporation).

RESULTS

A total of 187 subjects enrolled in this study: 122 with BCC and 65 with SCC. 171 subjects (91.4%) completed questionnaires at all three time points; 16 patients (8.6%) were lost to follow up (Figure 1). The mean age of respondents was 67 and 59% were male (Table 1). Tumors were more likely to be located on the head and neck (84%). Most patients had a previous history of skin cancer (63%).

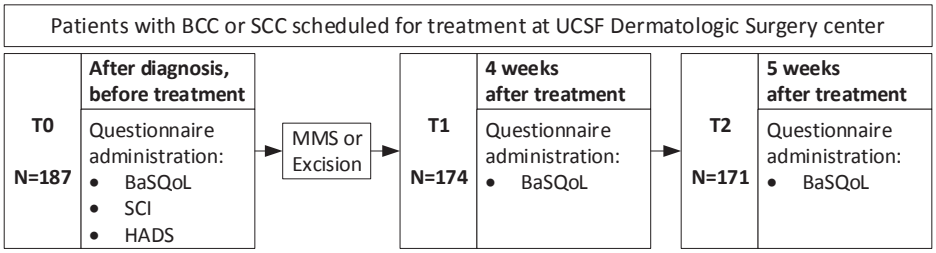

Figure 1. Study Design and Patient Flow Diagram. Patients with the diagnosis of SCC or BCC and scheduled for treatment in the UCSF Dermatologic Surgery center were approached for voluntary participation. Participants were asked to fill out BaSQoL, HADS, and SCI one week before their treatment (T0). They were then asked to complete BaSQoL four weeks after treatment (T1), and five weeks after treatment (T2).

Table 1. Patient characteristics

N participants T0	187
N participants T1	174
N participants T2	171
N= 187	**N (%)**
Gender	
• Male	111 (59)
• Female	75 (41)
Age mean (SD)	67 (13.9)

Table 1. Patient characteristics (*continued*)

Age groups	
• < 60	46 (25)
• 60-69	53 (28)
• 70-79	59 (32)
• ≥ 80	29 (16)
First skin cancer?	
• No	118 (63)
• Yes	49 (26)
• Unknown	20 (11)
Tumor site	
• 1	162 (87)
• multiple	25 (13)
Tumor site	
• scalp	14 (8)
• forehead	19 (10)
• nose	39 (21)
• eyelids	5 (3)
• cheeks	31 (17)
• lips	6 (3)
• other face	36 (19)
• neck	7 (4)
• trunk	14 (8)
• hands or feet	2 (1)
• extremity (not hands or feet)	14 (8)
Tumor type	
• BCC	122 (65)
- superficial	- 9 (5)
- nodular	- 71 (38)
- infiltrative	- 17 (9)
- morpheaform	- 3 (2)
- micronodulair	- 12 (6)
- infundibulocystic	- 2 (1)
- unknown	- 8 (4)
• SCC	65 (35)

Table 1. Patient characteristics (*continued*)

Treatment	
• Mohs surgery	170 (91)
• Conventional excision	17 (9)
Mohs rounds	
• 1	88 (47)
• 2	51 (27)
• 3	14 (8)
• ≥ 4	16 (9)

Overall scores for BaSQoL subscales were generally low, demonstrating a moderate impact on QoL. Patients with SCC tended to demonstrate higher levels of anxiety about their cancer than patients with BCC as measured by the Worries subscale (median score 0.9 [IQR: 0.5-1.2] for BCC and median score 1.0 [IQR 0.5-1.4] for SCC, p=0.013). The SCI similarly measured moderate overall impact on quality of life and higher impact on the Emotional subscale for SCC patients. HADS scores were almost uniformly low, except for a few patients who indicated anxiety. Table 2 summarizes scores for the various instruments before treatment (T0).

Table 2. HRQoL measurement T0

N = 187	BCC		SCC		p-value
BaSQoL (range 0-3)[#]	**N**	**Median score (IQR)**	**N**	**Median score (IQR)**	
• Behaviour	122	1.0 (0.5-1.3)	63	0.5 (0-1.3)	0.281
• Diagnosis & Treatment	122	1.0 (0.3-1.3)	65	1.0 (0.3-1.3)	0.987
• Worries	120	0.9 (0.5-1.2)	61	1.0 (0.5-1.4)	0.013
• Appearance	118	0.3 (0-1.0)	64	0.3 (0-1.0)	0.344
• Other People	120	1.0 (0.5-1.5)	64	0.5 (0.5-1.5)	0.913
SCI (range 0-100)[$]					
• Emotional	120	73.2 (57.1-82.1)	64	60.7 (39.3-81.3)	0.035
• Social	118	90.0 (80.0-95.0)	64	85.0 (71.3-98.8)	0.641
• Appearance	121	75.0 (50.0-92.0)	64	75.0 (50.0-100.0)	0.527
HADS (range 0-21)[#]					
• Anxiety	122	4.0 (2.0-7.0)	65	4.0 (2.0-7.0)	0.659
• Depression	121	1.0 (0-3.5)	63	2.0 (0-4.0)	0.865

[#] Higher score indicates higher impact on HRQoL

[$] Higher score indicates lower impact on HRQoL

Each item on the BaSQoL was tested using eight performance features (Table 3). Poor performance is defined as suboptimal performance on 3 or more features. None of the BaSQoL items met criteria for poor performance. Five out of 16 BaSQoL items showed one suboptimal feature, and three out of the 16 BaSQoL items showed 2 suboptimal performance features. Internal consistency was good with Cronbach's αs ranging from 0.63-0.80 for the different subscales (Behavior 0.80; Diagnosis & treatment 0.72; Worries 0.74; Appearance 0.76; Other people 0.63).

Table 3. Item performance of the BaSQoL questionnaire

N = 187	BaSQoL item number															
	Behavior				Other People	Diagnosis & treatment			Worries				Appearance			Other People
Item performance features	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Item difficulty																
Response distribution																
Item-test correlation			•		•						•					•
Item-rest correlation																
Item discriminant validity								•								
Item complexity											•					
Internal consistency					•											•
Stepwise regression		•								•					•	

• Indicates suboptimal performance in a given item feature. Definition of suboptimal performance in Supplementary Table 1.

The ICC between T1 and T2 was high (≥0.75) for nearly all subscales, indicating a stable response of the BaSQoL over time. Only the Worries subscale had moderate reliability (ICC 0.64, 95% CI: 0.53 – 0.72) (Table 4).

BaSQoL subscales were strongly correlated with subscales of the SCI, demonstrating convergent validity, but were weakly correlated with HADS, indicating divergent validity (Table 5).

We observed a trend towards improvement in 3 subscales (behavior, diagnosis & treatment and worries) and a trend towards deterioration in 2 subscales (appearance and other people) when comparing scores from before treatment (T0) to after treatment (T1). However, these changes did not reach statistical significance (Table 6).

Table 4. Test-retest reliability BaSQoL subscales T1 and T2
N = 171

Subscale	Mean (SD) score test	Mean (SD) score retest	Intraclass correlation coefficient*	95% CI	p-value
Behavior	0.87 (0.62)	0.83 (0.66)	0.795	0.731 – 0.845	<0.001
Diagnosis & Treatment	0.86 (0.63)	0.84 (0.62)	0.801	0.740 – 0.849	<0.001
Worries	0.80 (0.60)	0.75 (0.59)	0.635	0.533 – 0.719	<0.001
Appearance	0.55 (0.66)	0.52 (0.69)	0.745	0.668 – 0.807	<0.001
Other People	0.94 (0.70)	0.93 (0.70)	0.754	0.680 – 0.813	<0.001

* Two-way mixed effect model

Table 5. Convergent and divergent validity

N = 187	Multitrait-Multimethod correlation matrix using Spearman correlation coefficient				
	BaSQoL				
	Behavior	Diagnosis & Treatment	Worries	Appearance	Other People
Convergent validity: Skin Cancer Index					
Emotional	-0.247*	-0.532*	-0.721*	-0.369*	-0.361*
95% CI	-0.404/-0.091	-0.644/-0.402	-0.800/-0.623	-0.497/-0.231	-0.494/-0.213
Social	-0.810	-0.424*	-0.471*	-0.562*	-0.190*
95% CI	-0.244/0.078	-0.550/-0.282	-0.586/-0.340	-0.668/-0.445	-0.327/-0.046
Appearance	-0.201*	-0.441*	-0.288*	-0.670*	-0.134
95% CI	-0.372/-0.059	-0.569/-0.309	-0.424/-0.131	-0.757/-0.573	-0.279/0.013
Divergent validity: Hospital Anxiety and Depression Scale					
Anxiety	0.163*	0.402*	0.355*	0.394*	0.057
95% CI	0.020/0.315	0.260/0.525	0.205/0.492	0.247/0.518	-0.094/0.214
Depression	0.272*	0.235*	0.282*	0.376*	-0.020
95% CI	0.119/0.409	0.079/0.369	0.134/0.413	0.239/0.504	-0.177/0.131

* Correlation is significant at the 0.042 level (2-tailed) (FDR corrected).

■ are hypothesized high correlations

Table 6. BaSQoL subscale scores before and (4 weeks) after treatment

N = 174	Mean (SD)		Mean difference*	95% CI of the difference	p-value
BaSQoL subscale	T0	T1			
• Behavior	0.91 (0.71)	0.87 (0.62)	0.04	-0.05 – 0.13	0.44
• Diagnosis & Treatment	0.94 (0.67)	0.87 (0.65)	0.07	-0.21 – 0.15	0.14
• Worries	0.88 (0.58)	0.82 (0.61)	0.06	-0.02 – 0.14	0.16
• Appearance	0.53 (0.59)	0.58 (0.67)	-0.05	-0.15 – 0.05	0.34
• Other People	0.90 (0.71)	0.93 (0.71)	-0.04	-0.13 – 0.06	0.48

* Paired samples T-test

DISCUSSION

Health-related QoL has emerged as an essential outcome measure in dermatology. Therefore, disease-specific tools for measuring the impact of KC on QoL are needed. The BaSQoL questionnaire was developed following European Organisation for Research and Treatment of Cancer Quality of Life (EORTC QOL) group guidelines to comprehensively measure problems specific to patients with KC, such as fear of scars, coping mechanisms, worries about subsequent skin tumors, and the burden of sun protective behavior.

We have validated the English language translation of BaSQoL and demonstrated its utility in assessing an English-speaking American population. Using classical test theory, we have shown that BaSQoL performs well on all test features. Cronbach's αs were reasonable to high and demonstrated good internal consistency.

As hypothesized, BaSQoL correlated very well to SCI in our study population. However, BaSQoL also has several advantages over existing quality of life tools for KC. BaSQoL measures sun protective behavioral changes due to skin cancer, measures worries about treatment, and measures the QoL impact reliably over time. Three different sections of the questionnaire measure QoL impact since diagnosis, between diagnosis and treatment, and during the past week. This division of the questionnaire allows measurement of changing patient perceptions over time. In addition, BaSQoL addresses the impact of behavior change on patient QoL. The daily need for sun protection can be quite bothersome to patients with a new diagnosis of skin cancer, but is also one of the most important interventions to prevent further keratinocyte carcinomas. Measuring the impact of sun protection on patient QoL may give us insight into best practices to encourage behavior change.

Health-related QoL scores measured by both BaSQoL and SCI in this study indicated a modest impact compared to prior studies. [3,8] Our study population may have reported lower impact on QoL because the majority had a prior history of KC and may have been inured to the diagnosis. At least 63% had been treated for skin cancer previously and all of them were currently being treated at a tertiary referral center. Despite these low pre-treatment scores, the BaSQoL questionnaire was still able to detect a trend toward improvement in scores on 3 subscales (behavior, diagnosis & treatment and worries) and deterioration in 2 subscales (appearance and other people) after treatment. This demonstrates that this questionnaire is sensitive to change and could be used to compare treatment modalities for KC. Although the changes detected in this study did not reach statistical significance, as mentioned before, our study population was drawn from a tertiary referral center, had more experience with skin cancer and thus may have been less impacted by this diagnosis,

all of which would bias scores towards a smaller impact. Another reason for the small difference in BaSQoL scores before and after treatment may be the short time span between T0 and T1. Surgical scars may still be healing at 4 weeks postoperatively. Longer follow up may be needed to detect differences in HRQoL as measured by BaSQoL.

Strengths of this study include the comparison of the BaSQoL to two other HRQoL instruments, measurement of the BaSQoL at different time points, a sufficiently large sample of both BCC and SCC patients to test validity in both types of KC and excellent study completion with 91% of respondents completing the entire study. Limitations of this study include lack of a formal sample size calculation to measure sensitivity to change. As this is the first study measuring the BaSQoL before and after treatment, the anticipated difference was unknown before the start of the study and thus a formal calculation was not possible. The observed differences in the mean scores before and after treatment were very small, requiring at least more than 1,000 patients to show a statistical significant difference. As indicated before, larger differences may be observed within other patient populations or if longer periods of time are allowed between BaSQoL measurements.

In summary, the English language version of the BaSQoL has good face, content, and construct validity given its ability to measure moderate quality of life decrement in patients with keratinocyte carcinomas, its broad range of content drawn from patients and subject experts, and its convergence with SCI, and divergence with HADS. This study validates BaSQoL as a QoL measure for BCC and SCC patients over time. BaSQoL may be a useful tool in future studies to compare treatment modalities, interventions for sun protective behavior, or to identify patients in clinical practice with a substantial impact on their HRQoL who may benefit from additional clinical attention.

REFERENCES

1. Chren M-M, Sahay AP, Bertenthal DS, Sen S, Landefeld CS. Quality-of-life outcomes of treatments for cutaneous basal cell carcinoma and squamous cell carcinoma. *J Invest Dermatol*. 2007;127:1351-1357. doi:10.1038/sj.jid.5700740.
2. Lee EH. Patient expectations and performance measures in dermatologic surgery. *Clin Dermatol*. 2016;34(1):111-113. doi:10.1016/j.clindermatol.2015.07.002.
3. Rhee JS, Matthews BA, Neuburg M, Smith TL, Burzynski ML, Nattinger AB. Skin cancer and quality of life: assessment with the Dermatology Life Quality Index. *Dermatol Surg*. 2004;30(4 Pt 1):525-529. doi:10.1111/j.1524-4725.2004.30169.x.
4. Lee EH, Klassen AF, Nehal KS, Cano SJ, Waters J, Pusic AL. A systematic review of patient-reported outcome instruments of nonmelanoma skin cancer in the dermatologic population. *J Am Acad Dermatol*. 2013;69(2). doi:10.1016/j.jaad.2012.09.017.
5. Lee K, Weinstock M. Prospective quality of life impact of keratinocyte carcinomas: Observations from the Veterans Affairs Topical Tretinoin Chemoprevention Trial. *JAAD*. 2010;63(6):1107–1109.
6. Rhee JS, Matthews BA, Neuburg M, Logan BR, Burzynski M, Nattinger AB. The skin cancer index: clinical responsiveness and predictors of quality of life. *Laryngoscope*. 2007;117(3):399-405. doi:10.1097/MLG.0b013e31802e2d88.
7. Burdon-Jones D, Gibbons K. The Skin Cancer Quality of Life Impact Tool (SCQOLIT): A validated health-related quality of life questionnaire for non-metastatic skin cancers. *J Eur Acad Dermatology Venereol*. 2013;27(9):1109-1113. doi:10.1111/j.1468-3083.2012.04669.x.
8. Waalboer-Spuij R, Nijsten T. A review on quality of life in keratinocyte carcinoma patients. *G Ital Dermatol Venereol*. 2013;148:249-254.
9. Gibbons E, Casañas I Comabella C, Fitzpatrick R. A structured review of patient-reported outcome measures for patients with skin cancer, 2013. *Br J Dermatol*. 2013. doi:10.1111/bjd.12310.
10. Waalboer-Spuij R, Hollestein LM, Timman R, van de Poll-Franse L V, Nijsten TE. Development and Validation of the Basal and Squamous Cell Carcinoma Quality of Life (BaSQoL) Questionnaire. *Acta Derm Venereol*. 2018;98(2):234-239. doi:10.2340/00015555-2806.
11. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009. doi:10.1016/j.jbi.2008.08.010.
12. Sprangers MAG, Cull A, Groenvold M, Bjordal K, Blazeby J, Aaronson NK. The European Organization for Research and Treatment of Cancer approach to developing questionnaire modules: An update and overview. *Qual Life Res*. 1998. doi:10.1023/A:1008890401133.
13. Bullinger M, Alonso J, Apolone G, et al. Translating health status questionnaires and evaluating their quality: The IQOLA Project approach. *J Clin Epidemiol*. 1998. doi:10.1016/S0895-4356(98)00082-1.
14. Glickman ME, Rao SR, Schultz MR. False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *J Clin Epidemiol*. 2014. doi:10.1016/j.jclinepi.2014.03.012.

Supplementary Table 1. Definitions of item performance features used in classical test theory

Item	Item performance feature	Definition
1	Item difficulty	Proportion of missing scores among the 187 respondents. Item difficulty was considered high if 10% or more of scores were missing.
2	Response distribution	The proportion of patients who responded to each item with the same response was determined. An item was described as having a poor distribution if > 70% of patients had chosen the same response.
3	Item–test correlation	The Spearman's correlation coefficients (r) of each item with its subscale were calculated. If the r of an item differed > 0.1 with the r of the other items in the subscale, it was considered suboptimal.
4	Item–rest correlation	The Spearman's correlation coefficients (r) of each item with the sum of the other items in that subscale were calculated. Suboptimal item–rest correlation was defined as $r < 0.20$
5	Item discriminant validity	We compared the item–rest correlation coefficients with the correlation coefficients of an item with the other subscales. If the former equalled or was smaller than the latter, an item was defined as having poor discriminant validity.
6	Item complexity	We investigated the factor loadings in a factor analysis for each item. Suboptimal complexity was said to exist if the highest loading of an item was < 0.40 or if the difference between the loadings on different factors was < 0.10.
7	Internal consistency	For each subscale, the Cronbach's α was calculated. If $\alpha < 0.70$, the internal consistency was considered suboptimal for each subscale's item.
8	Stepwise regression	For each subscale, a forward stepwise regression analysis was performed. If an item entered the model after 90% or more of the variance of that subscale was explained it was considered suboptimal.