Development and validation of the basal and squamous cell carcinoma quality of life (BaSQoL) questionnaire

Rick Waalboer-Spuij
Loes M. Hollestein
Reinier Timman
Lonneke V. van de Poll-Franse
Tamar E.C. Nijsten

on behalf of the BaSQoL Group
Acta Derm Venereol. 2018 Feb 7;98(2):234-239.

ABSTRACT

Health-related quality of life (HRQoL) is important in basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) management. Disease-specific questionnaires exist, but with important shortcomings. Our goal was to develop and validate a questionnaire suitable for use in all BCC and SCC patients. In a four-phase trajectory, a preliminary questionnaire was created and tested population-based (1173 patients). The questionnaire was reduced using exploratory factor analysis and item response theory. Individual item performance was assessed using classical test theory. 721 patients completed the questionnaire. The number of items was reduced to 16, covering five scales. Confirmatory factor analysis showed a good fit. Cronbach's αs (range 0.67 – 0.82) were reasonable to high with good internal consistency. The Basal and Squamous cell carcinoma Quality of Life questionnaire has good face, content and construct validity. It is useful in the wide range of BCC and SCC patients and captures HRQoL impact in different timeframes.

INTRODUCTION

The use of patient-reported outcome measures (PROMs) and more specifically health-related quality of life (HRQoL) in dermatology patients has dramatically increased over the past decades. It is now an essential outcome for clinical studies and in daily practice, especially in chronic inflammatory skin diseases [1, 2]. In skin cancer, the use of PROMs and HRQoL has only been used over the past two decades and most of the focus has been on melanoma [3]. Since the incidence of basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) is increasing rapidly [4-6], the need for PROMs assessment including HRQoL is warranted to evaluate individual and global disease burden. Generic, cancer or even melanoma specific HRQoL instruments are neither content specific nor sensitive enough to detect the impact of the rarely life-threatening BCCs and SCCs that are most often treated by conventional excision. A specific issue for keratinocytic cancers is that patients are likely to develop multiple carcinomas and also actinic keratosis (AK) (so called actinic neoplasia syndrome) and that they can check their skin constantly [7].

Measurement of HRQoL in BCC patients has occurred in several studies, using generic, cancer related and dermatology specific questionnaires, all reporting little to no impact [7-13].

A few disease specific questionnaires have been developed, but these have several important shortcomings. The Skin Cancer Index (SCI) is developed and tested only in a tertiary care Mohs surgery clinic and therefore only suitable for use in a selected population [14, 15]. The Skin Cancer Quality of Life Impact Tool (SCQOLIT) is developed as a tool for non-metastatic skin cancer patients [16]. A limitation of the SCQOLIT is the addressing of five psychological issues regarding two different aspects in one item. In contrast to the SCI and the SCQOLIT , the Skin Cancer Quality of Life Questionnaire (SCQoL) was developed and validated with modern test theory, namely Rasch analysis [17]. This instrument was however derived from the previously developed Actinic Keratosis Quality of Life questionnaire (AKQoL) and pre-tested in a small sample (18 AK patients, 14 skin cancer patients) with the objective to distinguish the difference between AK and skin cancer patients [18]. From a content validity perspective, we feel that the before mentioned questionnaires do not to capture the psychological issues due to the often required behavioural changes to reduce sun exposure [19].

The objective of this study is to create and validate a HRQoL questionnaire suitable for BCC and SCC patients addressing relevant issues for patients and healthcare providers using different methodological approaches.

METHODS

Study design

The BCC and SCC specific HRQoL questionnaire was prepared and developed following the European Organisation for Research and Treatment of Cancer (EORTC) QOL group guidelines as much as possible [20-22]. However, the questionnaire is not an EORTC QOL group product and was not developed internationally. The development was conducted in four phases.

Phase I:

The main goal of phase I was to generate an extensive list of HRQoL issues relevant to BCC and SCC patients. One focus group meeting to discuss and generate HRQoL issues was facilitated by two independent psychologists with no in-depth skin cancer knowledge. The group consisted of 10 BCC and/or SCC patients with different types and numbers of tumours, treatments, gender and age. The audio recording of the focus group was analysed by RWS to extract as much issues as possible without formal transcribing. Extensive literature searches through PubMed (Table S4) and semi-structured interviews with 5 healthcare providers (HCP) provided additional issues [23].

The issues were discussed in an expert panel including dermatologists, psychologists and epidemiologists to identify the relevant disease-specific domains and issues (figure 1).

The remaining issues were presented to HCP (dermatologists, plastic surgeons, ophthalmologist, head-neck ENT surgeon, general practitioners) and patients for feedback and cognitive debriefing. They were also asked to rate the issues for relevance from 1 (not relevant) to 4 (very relevant) on a Likert scale (relevance rating). Issues with relevance mean score ≥ 1.5 were selected for priority rating. HCP and patients were asked to select 15 core issues to be included in the questionnaire (priority rating). Priority ratings of ≥ 30% were scored in het HCP group and ≥ 20% in the patient group. Issues scoring ≥ 3 criteria were included in the final issue list [20].

Phase II:

The final issue list was rephrased into questions compatible with the EORTC QLQ-C30 in terms of format of response categories [24]. The time frame of the questions was divided into three parts (‘since diagnosis’, ‘time between diagnosis and treatment’ and ‘during the past week’) since the items fitted different timeframes.

Phase III:

The item questionnaire was pretested in 16 patients.

Phase IV:

The questionnaire was field tested in 1,173 patients selected from the Netherlands Cancer Registry, as collected by Comprehensive Cancer Centre Netherlands, location Eindhoven. Patients were selected if they were diagnosed in one of the nine participating hospitals or clinics during the past twelve months before the field testing. The aim of the field testing was to determine scale structure, reliability, validity and to reduce the number of items. The Skindex-17 and the QLQ-C30 were also administered.

Statistical analysis

Descriptive statistics (means and percentages) were used in phase II to calculate relevance and priority ratings of the issue list and in phase IV to describe the patient characteristics. Type of BCC was grouped as multifocal (8091 of the International Classification of Disease for Oncology [ICD-O3]), infiltrating (8092, nodular (8097), other (8090,8093,8094,8095). Aforementioned analyses were performed in IBM SPSS Statistics for Windows, Version 21.0 (Armonk, New York: IBM Corporation).

After phase IV, the components were determined using principal component analyses (PCA) with varimax rotation. The number of components was determined with a Monte Carlo PCA for parallel analysis [25]. We ran two PCAs, one with complete cases and one with mean substitution, with one missing at most. Items with loadings of >0.40, were selected for Item Response Theory (IRT) [26]. IRT was used to select a minimum number of the best discriminating items covering the whole range of the latent traits.

For IRT analysis, we applied the two parameter latent trait model (2PL-ltm) [27] of the ltm package in R version 3.0.0. The 2PL-ltm program results in an ordering of the items on a given trait or component and supplies a discrimination value for each item. The 2PL-ltm program needs binary items as input. By collapsing the four answer category to binary items, some loss of information is induced. This method is preferred over multicategory models, because these do not provide an ordering of the items.

The original categories were “not at all”, “a little”, “quite a bit “ and “very much”. For the majority of items the median was between the first and second category, and for this reason we dichotomized between “not at all” and “a little” or more.

The items were selected on basis of their position on the relevant trait or component and their discriminative value. As we postulated an absolute maximum of five items per subscale, we divided the range between the lowest and highest position by five, and we choose from each of these intervals the item with the highest discriminative value. We

checked the unidimensionality of the remaining items with the “unidim” test of the ltm package.

After the item reduction by the 2PL-ltm model, item performance features as used in Classical Test Theory (CTT) were tested. The definitions of the features are presented in Table S5 [28, 29]. 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 also to test item discriminant validity. Internal consistency was tested via Cronbach’s α coefficients. Stepwise regression was performed 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.

The resulting factors were also tested with oblique confirmatory factor analyses. We applied two analyses, a complete cases analysis and a maximum likelihood analysis with missing values. We evaluated the fit indices according to the recommendations of Kline, Hu & Bentler and Brown[30-32]. The correlations between the subscales were reported. The confirmatory factor analyses were performed with STATA version 14.1 (College Station, Texas 77845 USA). All P-values were two sided and considered significant if $\alpha < 0.05$.

RESULTS

Phase I-IV:

The focus group meeting resulted in 63 issues, that were extended to 108 issues after literature searches (Table S4) and HCP interviews (Figure 1). After an expert consensus meeting 51 issues were eliminated from the list due to overlap of the issues, questions concerning information about the disease, cancer generic issues or other problems that were considered outside of the domain of HRQoL.

The remaining 57 issues were rated (mean scores, range, relevance and priority rating) by 42 patients (mean age of 70 years, 1-30 years since diagnosis, 27 BCC, 5 SCC and 10 diagnosis unknown to the patient) and 15 HCP (7 dermatologists, 1 plastic surgeon, 1 head neck ENT-surgeon, 1 ophthalmologist, 1 radiation oncologist and 4 general practitioners) and resulted in the removal of the 24 issues with lowest relevance and priority ratings (Figure 1).

The remaining 33 issues were constructed into a provisional 33 item questionnaire (Table S1)

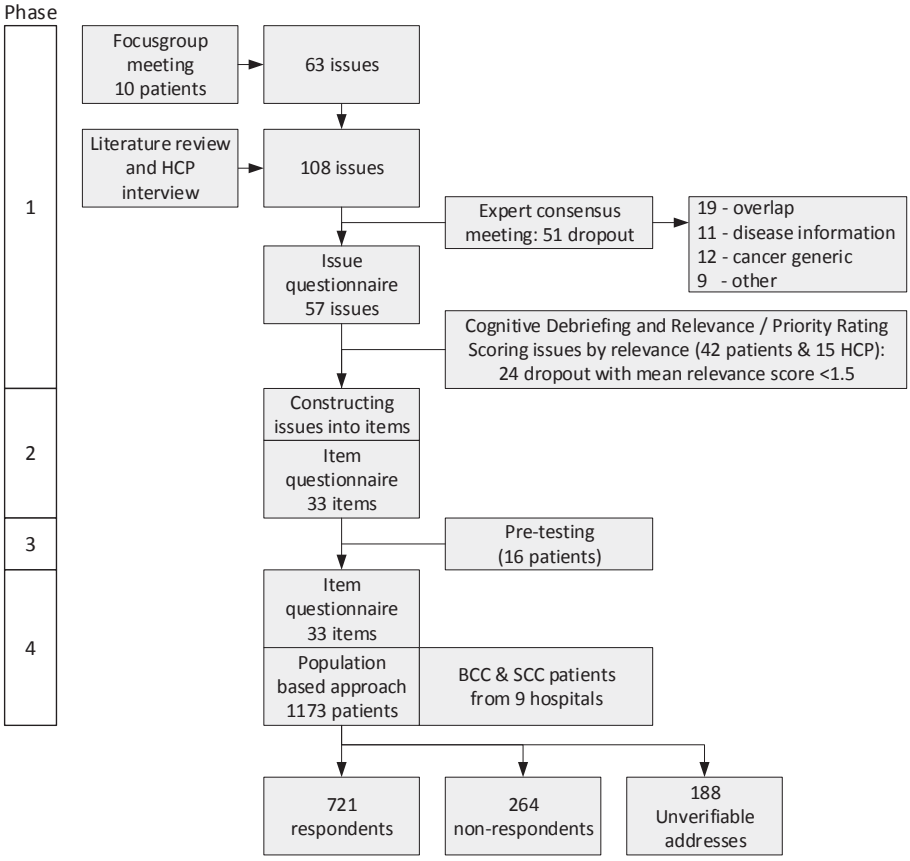

Figure 1. Questionnaire development phases.
HCP = Health Care Professional, BCC = Basal Cell Carcinoma, SCC = Squamous Cell Carcinoma.
Phase number as described by the EORTC QOL group guidelines.

This provisional questionnaire was reviewed by 16 patients for readability, clarity of the items and overlapping of the items and none of the items were excluded or rephrased.

The field testing was performed by selecting 1,173 BCC and SCC patients from 9 hospitals. The response rate was 61% and 721 patients completed the questionnaire. (Table 1) Of all respondents 85% had BCC and 15% had SCC.

The data contained 582 complete cases, 63 cases with one missing value and 76 cases with more missing values.

Table 1. Patient Characteristics

	Respondents	Non-respondents	Unverifiable addresses	p-value
Total number of patients (N)	721	264	188	
Sex	(column%)	(%)	(%)	0.0063
Male	51	37	49	
Female	49	63	51	
Age				
Mean, SD	67.3, 11.8	71.4, 13.5	61.3, 15.1	<0.0001
Median, IQR	68, 15	74.5, 16	61.5, 22.5	
<39	1	2	9	<0.0001
40-49	8	7	16	
50-59	14	9	21	
60-69	31	18	22	
70-79	32	33	18	
80+	14	31	13	
SCC (%)	15	16	9	0.0560
Socioeconomic status				
Low	17	22	13	<0.0001
Intermediate	28	29	20	
High	29	31	18	
Institute	3	4	4	
Unknown	23	13	46	
Location of tumour				
Face	78	78	85	0.1000
Other	22	22	15	
Other skin tumours*				
Multiple BCC	16	19	11	0.1000
Multiple SCC	9	7	6	#
MM	0	0	1	#
Other	0	0	0	#
The following variable is only available for BCC				
BCC (N)	613	222	171	
Type BCC	(column%)	(%)	(%)	
Multifocal	11	8	9	0.070
Infiltrating	18	22	15	
Nodular	64	65	65	
Other	7	4	12	

* patients can have combinations

No statistical test performed due to low numbers

Principal component analyses

The two PCAs (complete cases and with one missing included) both resulted in six components, with the same items loading. Items 23 and 24 formed a separate component, and at face value these items are nearly identical. Leaving out one of them resulted in five components. Item 24 had a higher factor loading than item 23, for this reason item 23 was removed from the analyses. Only item 5 was not eligible, because it had a component loading lower than 0.40.

The five components were labelled as: Worries (8 items, α=0.87), Appearance (7 items, α=0.84), Behaviour (7 items, α=0.85), Diagnosis & Treatment (5 items, α=0.84) and Other people (4 items, α=0.79) (Table 2).

Table 2. Subscales and item characteristics.

	Missing values	Principal component loading	2PL-ltm solution		Selected BaSQoL items	Unidim p-value
			Position	Discrimination		
Worries					α = 0.82	0.0297
19	2	.764	0.013	2.609	10	●
17	1	.724	-0.404	3.056	9	●
25	0	.696	0.249	2.079	12	●
26	0	.665	-0.164	2.206		
21	2	.646	0.827	2.472	11	●
28	0	.630	-0.458	2.357		
18	1	.626	-0.219	2.247		
24	1	.482	-0.115	0.619		
10	2	.401	0.035	1.112		
Appearance					α = 0.71	0.6733
33	0	.787	1.239	5.025	15	●
31	1	.779	1.151	4.414		
29	0	.770	1.144	3.987		
22	2	.725	1.008	3.389	13	●
30	3	.661	1.253	3.06		
15	1	.580	*			
32	9	.459	1.981	2.251	14	●
Behaviour					α = 0.79	0.6931
9	0	.838	0.162	3.985	4	●
4	7	.763	0.212	2.357		
6	1	.760	0.028	2.479		
1	1	.748	-0.099	2.846	1	●
2	2	.741	0.296	2.79	2	●

Table 2. Subscales and item characteristics. (*continued*)

	Missing values	Principal component loading	2PL-ltm solution		Selected BaSQoL items	Unidim p-value
			Position	Discrimination		
3	5	.568	0.748	1.297	3	●
5	1	.349				
Diagnosis & Treatment					α = 0.78	0.7426
12	7	.797	0.34	5.472	7	●
14	2	.745	0.624	2.146		
13	1	.686	0.654	2.381		
11	1	.610	-0.17	1.955	6	●
16	2	.509	-0.942	2.288	8	●
Other People					α = 0.67	1.000
8	2	.809	-0.403	2.362	5	●
7	1	.790	-0.491	2.299		
27	2	.705	0.130	3.624	16	●
20	0	.572	0.048	2.017		

* Item 15 prevented the program to converge, this items also had a high loading on the treatment component (0.371). Preliminary questionnaire item numbers are displayed in the first column (Table S1).

Item response analyses

The position on the components and discrimination values resulting from the 2PL-ltm analyses are presented in Table 2. On basis of these values the item set was reduced from 32 to 16 items. The characteristic curves of the selected items are presented in Figure 2. The "Worries" and "Behaviour" subscales retained 4 items (αs 0.79-0.82), the "Appearance" and "Diagnosis & Treatment" subscales retained three items (αs = 0.71-0.78) and the "Other people" subscale retained 2 items (α = 0.67). The unidim p-value for the 4 selected items of "Worries" was significant (p=0.03), indicating that this subscale was not sufficiently unidimensional. This lack of unidimensionality was caused by item 21. However, the unidim p-value of all 9 items was 0.38 indicating that all 9 items (including 21) belonged to an unidimensional subscale. We decided to include the item in the final questionnaire because we considered it to be a conceptually important aspect and because of the marking of the scale of the highest position on the latent trait. Item 15 in the "Appearance" prevented the program to converge. Inspection of this item showed that it also loaded (0.37) on the "Diagnosis & Treatment" subscale, and thus violated the unidimensionality assumption. It was decided to delete this item from the analyses. After this the unidim test was insignificant for the subscales appearance, behaviour, diagnosis & treatment and other people, indicating that the unidim assumption has been met for these subscales.

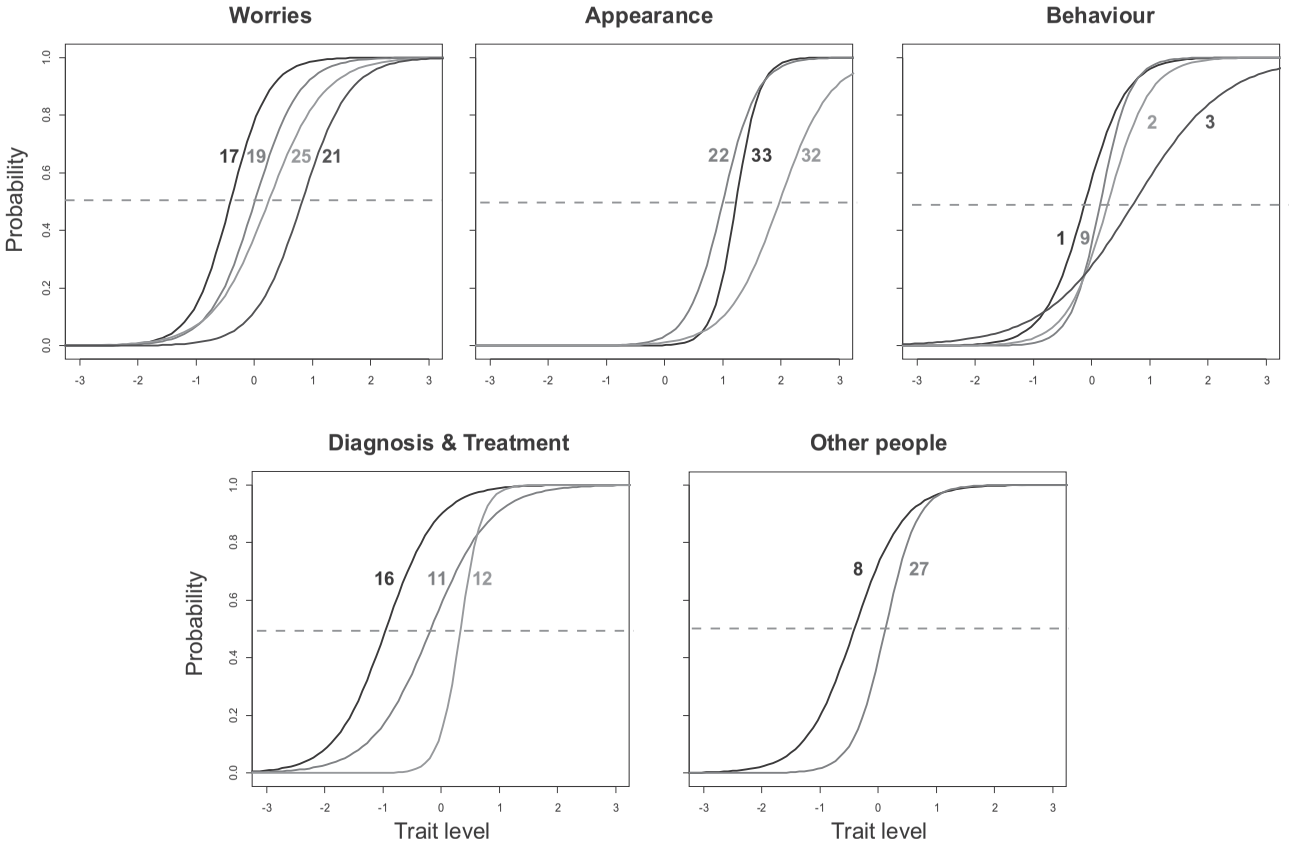

Figure 2. Item characteristic curves of the subscales.
The item characteristic curves depict the placement of the items on a latent ability and its discriminative value. For example, item 3 (provisional item number) discriminates best between patients with a high behavioural score, and item 1 discriminates best in patients with a low score. Additionally, item 9 discriminates better than item 3.

The resulting 16 item questionnaire was named Basal and Squamous cell carcinoma Quality of Life (BaSQoL) questionnaire. (Table S2)

Classical Test Theory

The eight CTT item performance features of the newly constructed questionnaire showed that 7 out of 16 items showed only one suboptimal feature and one showed two suboptimal performance features (Table 3). From a CTT perspective, the overall performance of the BaSQoL is therefore considered as good. There was no significant correlation with the subscales of the Skindex-17 and the QLQ-C30 suggesting different issues were captured.

Table 3. Item performance of the BaSQoL questionnaire

	BaSQoL subscales															
	Behaviour				Other People	Diagnosis & treatment			Worries				Appearance			Other People
BaSQoL item number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Item performance features																
Item difficulty																
Response distribution											•		•		•	
Item-test correlation																
Item-rest correlation																
Item discriminant validity																
Item complexity																
Internal consistency					•											•
Stepwise regression		•					•				•			•		
Provisional 33 item questionnaire number	1	2	3	9	8	11	12	16	17	19	21	25	22	32	33	27

• Indicates suboptimal performance in a given item feature. Definition of suboptimal performance in Table S5. Item numbers displayed are the final BaSQoL item numbers (Table S2)

Confirmatory Factor Analyses

Both the complete cases and the maximum likelihood with missing vales (MLMV) had acceptable to good misfit scores (RMSEA and SRMR) and good goodness of fit (CFI and TLI) (Table 4). The correlations between the subscales were generally low and there were only small differences between the two analyses. (Table S3).

Translation

The original Dutch version of the BaSQoL was translated into English by forward-backward translation [22]. (Table S1)

Table 4. Fit indices Confirmatory Factor Analysis

	Complete cases	MLMV	Recommended Kline	Recommended Hu & Bentler	Recommended Brown
Measures of misfit:					
RMSEA	0.050	0.053	<0.05	± 0.06	<0.05 / <0.08*
SRMR	0.042	--	<0.10	± 0.08	<0.08
Goodness of fit:					
CFI	0.958	0.956	>0.90		>0.95
TLI	0.947	0.944	>0.90	± 0.95	>0.95

MLMV - maximum likelihood with missing values
CFI - comparative fit index
SRMR - standardized root mean squared residual
RMSEA - root mean squared error of approximation
TLI - Tucker-Lewis index
* <0.05 – good, <0.08 reasonable

Scoring

The individual items are scored from 0 to 3, in which 0 represents no impact and 3 very high impact. The mean score per subscale is calculated as a scale score. A minimum of 50% of the questions within the subscale has to be answered to calculate the subscale score.

DISCUSSION

The BaSQoL questionnaire has been developed methodologically by following the EORTC QOL group guidelines as much as possible [20-22] and assesses the relevant dimensions of HRQoL in BCC and SCC patients.

The content of the BaSQoL questionnaire shows some overlap with items from the existing questionnaires for skin cancer, such as cancer recurrence or spreading, concerns about scarring and sun behaviour. But the BaSQoL captures a broader spectrum of the issues relevant in BCC and SCC patients such as treatment and diagnosis related issues and long-term behavioural changes [14, 16, 17]. Since our questionnaire was developed and validated in a large Dutch patient sample by using a population based approach, we consider it to be representative for use in the wide range of BCC and SCC patients.

Since patients were extensively involved in the whole process of the development, the questions are representative and in the terminology as used by the patients.

By combining the use of modern IRT and CTT analyses we aimed to create a questionnaire with optimal psychometric properties. Therefore the BaSQoL has good face, content and construct validity.

The use of the different time frames in our questionnaire is also a unique feature. Patients noted a difference in behaviour before and after the initial diagnosis. Therefore the impact of this behavioural change is measured in the first part of the BaSQoL. The second part of the BaSQoL concerns the period of diagnosis and treatment. This, usually short, timeframe has a high impact on patients HRQoL. This subscale is suitable for assessing the patient's experience of this specific period in order to manage anxiety in the process in case of new tumours and, in general, to optimize patient care. The final part of the questionnaire addresses the impact of the skin cancer during the past week. Since BCC and SCC are being considered as more chronic diseases, addressing the relevant issues at the right moment is important.

The preliminary validation of the BaSQoL has also been established by this study. Cronbach's αs of the reduced subscales remained reasonable, taking into account that a reduction in the number of items generally leads to lower α [33, 34]. The subscales are psychometrically robust, displaying excellent item performance and a good fit in the confirmatory factor analysis. As the BaSQoL measures different aspects of HRQoL, it showed no significant correlation with the subscales of the Skindex-17 and the QLQ-C30 confirming divergent validity. Unfortunately, none of the previously developed BCC or SCC specific questionnaires were included in this study because there are no validated BCC or SCC specific questionnaires available in Dutch and we intended to minimize respondent burden and increase the response rate. A validation study of the English version of the BaSQoL is underway. Construct validity by comparing to the validated SCI, test-retest stability and responsiveness to change will be addressed in this study. Other important features to increase interpretability such as categorization of scores and minimally clinical important difference remain to be determined.

Item 21 (BaSQoL nr 11) 'Were you uncertain about the future?', that violated the unidim assumption of the worries subscale, also had a suboptimal response distribution (Table 3). The confirmatory factor analysis however, showed a good fit. This item reflects a more generic aspect than the other items in the subscale and it had far the highest position on the latent trait for this reason and because of the conceptual general intent of the item we decided to maintain it within the questionnaire.

In summary, the BaSQoL has good face, content and construct validity. The BaSQoL is representative for use in the wide range of BCC and SCC patients and captures HRQoL impact in different time periods. Therefore we consider the BaSQoL as a useful tool to capture HRQoL impact in future studies.

ETHICAL CONSIDERATIONS

This study was approved by the local ethics committee of the Erasmus Medical Centre Rotterdam (Reference number MEC-2013-420)

CONFLICTS OF INTEREST AND FUNDING

The authors state no conflict of interest.

This study was financially supported by a grant by Roche, Leo Pharma and Galderma.

ACKNOWLEDGMENTS

Members of the BaSQoL Group: M.J. Aarts, M.T. Bastiaens, O. Husson, B.A. Jagtman, F.H.J. Koedijk, D.I.M. Kuijpers, H.C.J. Liberton, J.M. Mommers, S. Oerlemans, K.P. de Roos, M.W.H. Timmermans, E. de Vries, L.J.M.T. Weppner-Parren.

REFERENCES

1. Both H, Essink-Bot ML, Busschbach J, Nijsten T. Critical review of generic and dermatology-specific health-related quality of life instruments. J Invest Dermatol. 2007;127(12):2726-39.
2. Prinsen CA, de Korte J, Augustin M, Sampogna F, Salek SS, Basra MK, et al. Measurement of health-related quality of life in dermatological research and practice: outcome of the EADV Taskforce on Quality of Life. J Eur Acad Dermatol Venereol. 2013;27(10):1195-203.
3. Livingstone E, Krajewski C, Eigentler TK, Windemuth-Kieselbach C, Benson S, Elsenbruch S, et al. Prospective evaluation of follow-up in melanoma patients in Germany - results of a multicentre and longitudinal study. Eur J Cancer. 2015;51(5):653-67.
4. Flohil SC, de Vries E, Neumann HA, Coebergh JW, Nijsten T. Incidence, prevalence and future trends of primary basal cell carcinoma in the Netherlands. Acta Derm Venereol. 2011;91(1):24-30.
5. Flohil SCS, I.; van Rossum, M.M.; Coebergh, J.W.; de Vries, E.; Nijsten, T. Trends in Basal Cell Carcinoma Incidence Rates: A 37-Year Dutch Observational Study. J Invest Dermatol. 2012.
6. Hollestein LM, de Vries E, Nijsten T. Trends of cutaneous squamous cell carcinoma in the Netherlands: increased incidence rates, but stable relative survival and mortality 1989-2008. Eur J Cancer. 2012;48(13):2046-53.
7. Weinstock MA, Lee KC, Chren MM, Marcolivio K. Quality of life in the actinic neoplasia syndrome: The VA Topical Tretinoin Chemoprevention (VATTC) Trial. J Am Acad Dermatol. 2009;61(2):207-15.
8. Blackford S, Roberts D, Salek MS, Finlay A. Basal cell carcinomas cause little handicap. Qual Life Res. 1996;5(2):191-4.
9. Chren MM, Lasek RJ, Quinn LM, Covinsky KE. Convergent and discriminant validity of a generic and a disease-specific instrument to measure quality of life in patients with skin disease. J Invest Dermatol. 1997;108(1):103-7.
10. Roberts N, Czajkowska Z, Radiotis G, Korner A. Distress and Coping Strategies Among Patients with Skin Cancer. J Clin Psychol Med Settings. 2012.
11. Rhee JS, Matthews BA, Neuburg M, Smith TL, Burzynski M, Nattinger AB. Skin Cancer and Quality of Life: Assessment with the Dermatology Life Quality Index. Dermatol Surg. 2004;30(4 I):525-9.
12. Chren MM, 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(6):1351-7.
13. Sampogna F SA, Di Pietro C, Pagliarello C, Paradisi A, Tabolli S, Abeni D. Field performance of the Skindex-17 quality of life questionnaire: a comparison with the Skindex-29 in a large sample of dermatological outpatients. J Invest Dermatol. 2013;133(1):104-9.
14. Rhee JS, Matthews BA, Neuburg M, Burzynski M, Nattinger AB. Creation of a quality of life instrument for nonmelanoma skin cancer patients. Laryngoscope. 2005;115(7):1178-85.
15. 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.
16. 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 Dermatol Venereol. 2012.

17. Vinding GR, Christensen KB, Esmann S, Olesen AB, Jemec GB. Quality of life in non-melanoma skin cancer--the skin cancer quality of life (SCQoL) questionnaire. Dermatol Surg. 2013;39(12):1784-93.

18. Esmann S, Vinding, G.R., Christensen, K.B., Jemec, G.B.E. Assessing the influence of Actinic Keratosis on Patients' Quality of Life - The AKQoL Questionnaire. Br J Dermatol. 2013;168(2):277-83.

19. Waalboer-Spuij R, Nijsten TE. A review on quality of life in keratinocyte carcinoma patients. G Ital Dermatol Venereol. 2013;148(3):249-54.

20. Blazeby J, Sprangers MA, Cull A, Groenvold M, Bottomley A. EORTC quality of life group: Guidelines for developing questionnaire modules. Brussels: EORTC; 2002.

21. Sprangers MA, Cull A, Bjordal K, Groenvold M, Aaronson NK. The European Organization for Research and Treatment of Cancer. Approach to quality of life assessment: guidelines for developing questionnaire modules. EORTC Study Group on Quality of Life. Qual Life Res. 1993;2(4):287-95.

22. Sprangers MA, 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. EORTC Quality of Life Study Group. Qual Life Res. 1998;7(4):291-300.

23. Brod M, Tesler LE, Christensen TL. Qualitative research and content validity: developing best practices based on science and experience. Qual Life Res. 2009;18(9):1263-78.

24. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst. 1993;85(5):365-76.

25. Watkins MW. Determining Parallel Analysis Criteria. Journal of Modern Applied Statistical Methods. 2006;5(2):8.

26. Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment, and selection of questionnaires. Optometry & Vision Science. 2007;84(8):663-74.

27. Rizopoulos D. Package ltm: Latent trait models under IRT 2012 [Available from: http://rwiki.sciviews.org/doku.php?id=packages:cran:ltm].

28. Chren MM, Lasek RJ, Flocke SA, Zyzanski SJ. Improved discriminative and evaluative capability of a refined version of Skindex, a quality-of-life instrument for patients with skin diseases. Arch Dermatol. 1997;133(11):1433-40.

29. Streiner DL, Norman GR. Healt Measurements Scales: a Practical Guide to Their Development and Use. Oxford: Oxford University Press; 2003.

30. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal. 1999;6(1):1-55.

31. Kline RB. Principles and practices of structural equation modeling. New York: Guilford Press; 1998.

32. Brown TA. Confirmatory factor analysis for applied research. New York: The Guilford Pres; 2006.

33. Massof RW. The measurement of vision disability. Optometry & Vision Science. 2002;79(8):516-52.

34. Guilford JP. Psychometric Methods. New York, London: McGraw-Hill; 1936.

Table S1. Provisional item list

Number	Item	
1	Does it bother you to be careful about your behaviour in the sun?	●
2	Does it bother you to use more sunscreen (cream, spray, etc.)?	●
3	Does it bother you to check your skin for skin cancer?	●
4	Does it bother you to adjust your vacation to avoid the sun?	
5	Do you feel that you have to avoid direct sunlight?	
6	Does it bother you to wear certain clothing or a hat to protect you from the sun?	
7	Do you feel that you should warn others for the sun?	
8	Do you feel that you have to encourage others to get their skin checked?	●
9	Does it bother you to have to protect your skin from the sun?	●
10	Did you have the feeling having no control over your skin cancer?	
11	Were you worried about the period between diagnosis and treatment?	●
12	Were you afraid of the treatment?	●
13	Were you worried about (possible) side-effects of the treatment?	
14	Were you worried about the anaesthetic injections?	
15	Were you worried about scarring?	
16	Were you frightened by the word cancer?	●
17	Were you afraid to get skin cancer on multiple body sites?	●
18	Were you worried that the skin cancer would come back at the treated area?	
19	Were you worried about skin cancer spreading to other parts of the body?	●
20	Were you worried about family members getting skin cancer?	
21	Were you uncertain about the future?	●
22	Were you worried that you would be less attractive?	●
23	Was your skin itching at the skin cancer area?	
24	Was your skin sensitive at the skin cancer area?	
25	Were you worried about other skin disorders?	●
26	Were you insecure about not being able to recognise the signals of skin cancer?	
27	Were you worried about other people's skin?	●
28	Were you worried about the severity of skin cancer?	
29	Were you ashamed of the scar(s)?	
30	Did the questions by others about your scar(s) bother you?	
31	Were you worried about whether your scar(s) could be covered?	
32	Did it bother you to adjust your clothing in order to cover your scars and spots?	●
33	Did you feel less attractive?	●

● These items were translated from Dutch through forward/backward translating and are included in the final questionnaire.

Table S2. Basal and Squamous cell carcinoma Quality of Life questionnaire

The following questions are about the influence of skin cancer on your daily life

	Since the skin cancer diagnosis,	Not at all	A little	Quite a bit	Very much
1.	Does it bother you to be careful about your behaviour in the sun?	☐	☐	☐	☐
2.	Does it bother you to use more sunscreen (cream, spray, etc.)?	☐	☐	☐	☐
3.	Does it bother you to check your skin for skin cancer?	☐	☐	☐	☐
4.	Does it bother you to have to protect your skin from the sun?	☐	☐	☐	☐
5.	Do you feel that you have to encourage others to get their skin checked?	☐	☐	☐	☐
	When you think back to the time of diagnosis and treatment,	**Not at all**	**A little**	**Quite a bit**	**Very much**
6.	Were you worried about the period between diagnosis and treatment?	☐	☐	☐	☐
7.	Were you afraid of the treatment?	☐	☐	☐	☐
8.	Were you frightened by the word cancer?	☐	☐	☐	☐
	During the past week,	**Not at all**	**A little**	**Quite a bit**	**Very much**
9.	Were you afraid to get skin cancer on multiple body sites?	☐	☐	☐	☐
10.	Were you worried about skin cancer spreading to other parts of the body?	☐	☐	☐	☐
11.	Were you uncertain about the future?	☐	☐	☐	☐
12.	Were you worried about other skin disorders?	☐	☐	☐	☐
13.	Were you worried that you would be less attractive?	☐	☐	☐	☐
14.	Did it bother you to adjust your clothing in order to cover your scars and spots?	☐	☐	☐	☐
15.	Did you feel less attractive?	☐	☐	☐	☐
16.	Were you worried about other people's skin?	☐	☐	☐	☐
	Scoring of the items	**0**	**1**	**2**	**3**

Table S3. Correlations between subscales.

	Worries	Appearance	Behaviour	Diagnosis & Treatment	Other People
Worries		0.11	0.15	0.33	0.20
Appearance	0.14		0.08	0.13	0.04
Behaviour	0.15	0.09		0.16	0.09
Diagnosis & Treatment	0.33	0.16	0.15		0.17
Other People	0.20	0.06	0.09	0.17	

Lower triangle complete cases, upper triangle maximum likelihood with missing values (MLMV).

Table S4. Keywords used in PubMed searches

quality of life
health-related quality of life
basal cell carcinoma
squamous cell carcinoma
non-melanoma skin cancer

Table S5. Definitions of item performance features used in classical test theory

Item	Item performance feature	Definition
1	**Item difficulty**	Proportion of missing scores among the 721 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[a], 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[a] 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[a]. 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 α < 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.

[a] Subscales derived from the principal component analysis were used.