

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/burns

Convergent and discriminant validity of quality of life measures used in burn populations



Jill Meirte^{a,b,*}, Ulrike Van Daele^b, Koen Maertens^{a,c}, Peter Moortgat^a,
Rudi Deleus^d, Nancy E. Van Loey^{e,f}

^a Oscare, Organisation for Burns, Scar After-Care and Research, Antwerp, Belgium

^b University of Antwerp, Department of Rehabilitation Sciences and Physiotherapy, Antwerp, Belgium

^c Vrije Universiteit Brussel, Department of Clinical and Lifespan Psychology, Brussels, Belgium

^d UZ Gasthuisberg Leuven, Belgium

^e Association of Dutch Burn Centers, Department of Behavioural Research, Beverwijk, The Netherlands

^f Utrecht University, Department of Clinical Psychology, Utrecht, The Netherlands

ARTICLE INFO

Article history:

Accepted 14 July 2016

Keywords:

Patient reported outcome measures

Quality of life

Psychometrics

Burns

Functioning

ICF

ABSTRACT

Introduction: The primary goal of this study was to investigate convergent validity, i.e. overlapping constructs, of the Burn Specific Health Scale-Brief (BSHS-B), the Short Form-36 items Health Survey (SF-36) and the European Quality Of Life Five Dimensions (EQ-5D) within the International Classification of Functioning Disability and Health (ICF) framework. A secondary goal was to examine the discriminant validity of the questionnaires according to burn severity (no surgery versus 1 or more surgeries).

Methods: A prospective multi-centre study in adult patients with burns was conducted. At the 9 months assessment, two generic questionnaires, i.e. the SF-36 and EQ-5D, and the BSHS-B were completed. Pearson correlations were used to evaluate convergent validity. Linear discriminant analysis was used to evaluate discriminant validity.

Results: At 9 months post-burn data from 184 persons were available of which 131 (71%) were male, mean TBSA burned was 11.8% (SD = 10.2). Sixty five (34%) patients did not need surgery, 128 (66%) patients required one or more surgeries. Higher convergence was shown between the generic SF-36 and the condition specific BSHS-B whereas the EQ-5D showed lower convergence with the BSHS-B especially in the domain Activity. The generic scales discriminated across all scales whereas not all BSHS-B scales were able to differentiate problem levels across burn severity groups.

Conclusion: This study demonstrates that the ICF is useful to classify scales in order to identify overlapping areas as well as to uncover gaps in relation to patient reported outcomes. Both the SF-36 and EQ-5D showed the ability to distinguish levels of functioning across burn severity groups. As the BSHS-B performed less well and relevant domains of functioning were not addressed, there is room for improvement and modification of this condition specific questionnaire to better capture burn patients' functioning.

© 2016 Elsevier Ltd and ISBI. All rights reserved.

* Corresponding author at: University of Antwerp, Department of Rehabilitation Sciences and Physiotherapy, Antwerp, Belgium.
Tel.: +32 486887503.

E-mail address: jill.meirte@uantwerpen.be (J. Meirte).

<http://dx.doi.org/10.1016/j.burns.2016.07.001>

0305-4179/© 2016 Elsevier Ltd and ISBI. All rights reserved.

1. Introduction

During the last decade it has been acknowledged that functional outcome is an important parameter of treatment following burns [1]. Functional outcome is often measured using patient reported outcome measures (PROMs). Expert consensus exists on using both generic and disease-specific Quality Of Life (QOL) questionnaires to capture the full impact of a health condition [2]. However, it is insufficiently understood if this combination indeed comprises the full spectrum of functioning as defined by the International Classification of Functioning, Disability and Health (ICF). The ICF is a worldwide used framework to describe the health condition of a patient in a bio-pycho-social context and it can also act as a framework to inventory and compare the content of PROMs [3].

A previous theoretical classification showed the usefulness of the ICF to identify overlap and gaps in scales such as the Burn Specific Health Scale Brief (BSHS-B), the Short Form-36 items (SF-36) and the European Quality of Life 5 Dimensions (EQ-5D) [4]. In general the theoretical classification revealed that the subscales of the three questionnaires covered the Body function, Activity and Participation domains of the ICF. None of the subscales included items on Environmental factors and Body structures. The BSHS-B covered most ICF domains and was the only one to include Personal factors. The content comparison of the three questionnaires revealed considerable overlap across the questionnaires, particularly in the domain Body Function and Activity domains [4]. To the best of our knowledge, there are no studies that empirically evaluate the overlap between generic and disease specific scales within the ICF framework.

Furthermore, it is important to investigate if the scales are able to discriminate between burn severity groups. Discriminant validity is a characteristic of a measure that evaluates the ability to discriminate between groups with known differences. Prior studies [5,6] used total body surface area (TBSA) affected, number of surgeries, or length of stay (LOS) in hospital as burn severity measures. The reliability and validity of the BSHS-B was recently examined in minor burns [7] but it remains unclear how well the different subscales can discriminate between burn severity groups. However, the EQ-5D was sensitive for different levels of burn severity [8]. In a prospective longitudinal study LOS was associated with more problems in all EQ-5Ds dimensions excluding Usual Activities at 3 and 6 months. At 12 months post burn longer LOS was still associated with more problems in Mobility and Pain [8]. The number of surgeries [9] has been shown to negatively affect health related quality of life (HRQOL) [6] and was suggested to outperform somewhat better compared to TBSA as a predictor of long-term HRQOL and speed of recovery [6]. As subscales of PROMs are more often used to identify problem areas in specific health domains [10], it is interesting to evaluate whether the scales are able to detect differences in problem levels in groups that are known to differ from each other.

The aim of this study was twofold: (1) to empirically investigate the overlapping constructs of the BSHS-B, the EQ-5D and the SF-36 within the ICF framework. It was hypothesized that overlapping subscales should show correlations

exceeding .50 [11,12] and (2) to examine discriminant validity, i.e., how well the scales are able to discriminate between burn severity groups.

2. Methods

2.1. Patients

This study was reviewed and approved by the ethics committees of the Martini Hospital, Groningen, The Netherlands, and Ghent University Hospital, Belgium. A prospective multicenter cohort study using the BSHS-B, the SF-36 and EQ-5D to assess QOL 9 months after burn was conducted. The results of this study are part of a larger prospective cohort study on psychological problems and quality of life [6,13]. Adult patients with acute burns admitted to one of six burn centres in the Netherlands and Belgium between March 2003 and April 2005 requiring hospitalization for at least 72 h were included in this study. Patients were excluded when they were unable to complete the self-report questionnaires because of poor Dutch proficiency or cognitive disorders. In this study 311 adult burn patients met the inclusion criteria. Fifty-one patients (16%) refused to participate. Informed consent was obtained from 260 patients. We obtained 9 months follow-up data from 184 patients (71%).

2.2. Procedure

Eligible patients were invited to participate in the study during hospitalization. After signing the consent form the first questionnaires were administered during hospitalization. Once patients left the burn centre follow-up questionnaires were sent by regular mail, including a letter and a prepaid return envelope. Non-responders were reminded within 1 month after the questionnaires were sent. No further efforts were undertaken to collect the measures.

2.3. Measures

The BSHS-B, EQ-5D and SF-36 were completed 9 months after the burn. The EQ-5D is a widely used generic assessment tool which distinguishes five dimensions: Mobility, Self-Care, Usual Activities, Pain/Discomfort and Anxiety/Depression [8,6,14]. The EQ-5D covers three domains of the ICF; the domain (impairments in) Body function with subscales Pain/Discomfort and Anxiety/Depression, the Mobility and Self-Care subscales assess Activity (limitations) and Usual Activities inventories Participation (restrictions) [4]. Each dimension or subscale has three response levels: 1 = no problems, 2 = moderate problems or 3 = extreme problems. A Visual Analogue Scale for General Health can also be part of the EQ-5D but was not included in the analyses of this study. The EQ-5D was translated into more than 60 languages and is used worldwide [8]. The EQ-5D is short and easy to use.

The SF-36 is a generic measure made up of 36 items that are combined to form eight subscales [15,16] and covers (impairments in) Body Function with its subscales Mental Health, Vitality and Bodily Pain. Within the ICF domain Activity (limitations) the subscale Physical Functioning was found and

the Participation (restriction) domain is covered by multiple subscales: Role-Physical, part of the subscale Bodily Pain, Social Functioning and Role-Emotional [4]. Scoring is a two-step process. Most (29) items responses comprise Likert-type scales (three or six point scales), seven items have a binary response set (1 = 0 or 2 = 100). Each of the question responses relate to a different pre-coded numeric value. For each of the eight subscales an aggregate percentage score is produced. The percentage scores range from 0% (lowest or worst possible health) to 100% (highest or best health). In the second step scores are averaged across items constituting a scale [17,18].

The BSHS-Brief (BSHS-B) is the most recent version of this burn specific instrument. It includes 40 items comprising 9 domains [19,20], covering each domain of the ICF framework. Heat sensitivity, Sexuality and Affect comprised Body Function, Hand Function and Simple Abilities measure Activity limitations, Work and Interpersonal Relationship cover the domain Participation and Treatment Regimens and Body Image assess Contextual Factors [4]. Item responses are scored on a 5-point scale ranging from 0 (=all the time/great difficulty) to 4 (=never/no difficulty). Mean scores are calculated for each subscale and high scores indicate a good perceived health status [4].

2.4. Statistical analysis

Construct validity was examined using a priori hypotheses for discriminant and convergent properties of the instrument. Convergent validity refers to how well subscales correlate with other measures that are assumed to be related. Convergent validity, assessing the overlap between measures, was investigated using Pearson correlation coefficients among different subscales [11,21]. The degree of correlations

is outlined as follows: small 0.10 to 0.29, moderate ≥ 0.30 to 0.49 and strong ≥ 0.50 [22]. We hypothesized that subscales of the EQ-5D, the SF-36 and the BSHS-B theoretically classified as overlapping, were highly correlated (Table 1). Strong correlations indicate that the subscales assess similar constructs [11].

Discriminant validity evaluates the ability to discriminate between groups with known differences [23]. In this study discriminant validity was assessed by comparing two groups: those requiring no surgery versus those requiring ≥ 1 surgery. To describe discriminant validity a linear discriminant analysis was used to determine classification capacity of the different subscales of the SF-36, EQ-5D and BSHS-B in the two predefined groups (no surgery versus ≥ 1 surgery). In the analysis Box's M analysis for evaluation of equality of covariances and calculation of correctly predicted classification or hit ratio were included. Statistical significance was considered if $p < .05$. Analysis was performed using SPSS 20.0.

3. Results

3.1. Demographic and clinical characteristics of patients

A total of 184 participants completed the 9 month assessment of which 131 (71%) were male and 53 (29%) were female. Participants were on average 39.0 years old (SD = 12.8). The mean TBSA burned was 11.8% (SD = 10.2). The mean LOS in hospital was 23.0 days (SD = 21.0). The median number of surgeries was 1 (ranging between 0 and 16, SD = 1.8). Sixty five (34%) patients did not need surgery, 128 (66%) patients required one or more surgeries.

Table 1 – Coverage of ICF domains within BSHS-B, SF-36 and EQ-5D.

	Body function/structures (impairments)		Activity (limitations)	Participation Restrictions	Contextual factors		Others
	body structures	body functions	Activity	Participation	Personal factors	Environmental factors	
BSHS-B subscales		Heat sensitivity	Hand function	Interpersonal relationships~	Treatment regimes		
		Sexuality	Simple abilities***	Work~	Body image		
		Affect*					
SF-36 subscales		Mental health*	Physical functioning***	Role physical~			General health
		Bodily pain**		Social functioning~			
		Vitality		Role emotional~			
EQ-5D subscales		Pain/discomfort**	Mobility***	Usual abilities~			General health
		Anxiety/depression*	Self care***				

bold = unique features for the questionnaire.

italic = overlap between questionnaires within ICF domains.

*and arrows = illustrating link with other QOL scale.

3.2. Convergent validity between BSHS-B, EQ-5D and SF-36 subscales

Correlations between the subscales of the BSHS-B and the EQ-5D are shown in the upper part of Table 2. The negative correlations reflect a high score on an EQ-5D subscale as indicative of poor health state while a high score on the BSHS-B subscale is indicative of better health. Only two correlations exceeded .50. Within the Body Function domain, Affect ($r = -0.69$) was most strongly related with Anxiety/Depression. Within the Participation domain, Work (BSHS-B) and Usual Abilities (EQ-5D) revealed a strong ($r = -0.63$) correlation. Correlations within the Activity domain across the questionnaires were low.

Correlations between the subscales of BSHS-B and the SF-36 as categorized within the ICF domains are shown in the lower part of Table 2. High scores on the SF-36 and BSHS-B reflect a good health state. Again, within Body function, the Affect subscales were strongly related ($r = -0.67$). Within the Participation domain several high correlations were observed, with Work and Role Physical showing the highest correlation. Within the Activity domain, Physical Functioning and Hand Function were related. When comparing the overlap between

the two generic scales, the theoretically assumed overlap was confirmed by correlations exceeding 0.50. Only one correlation was lower, this being Role Emotional and Usual Activities. Correlations are shown in Table 3.

The highest correlations were indeed observed between subscales that were classified within the same ICF domain. However, other significant correlations appeared, providing insight in how domains are inter-connected. A combination of the scales provides additional insight in problem areas.

3.3. Discriminant validity of the HRQOL subscales between burn severity groups

Table 4 presents the mean scores for the BSHS-B for the two severity groups. A higher problem level was found in the ≥ 1 -surgery group. However the differences between the severity groups in the subscales Affect, Heat Sensitivity, Simple Abilities, Hand Function and Interpersonal Relationship did not reach statistical significance. The largest differences between groups were found with respect to Work, Body Image and Treatment Regimens. When tested by Box's M the subscale Work showed a significant unequal co-variance between both groups and should be interpreted with caution.

Table 2 – Pearson correlations between subscales of the BSHS-B the EQ-5D and the SF-36.

ICF domain	ICF domain	Body function			Activity		Participation		Contextual factors	
	BSHS-B	Affect	Sexuality	Heat Sensitivity	Simple Abilities	Hand Function	Work	Inter. Relat.	Body Image	Treatment Regimens
ICF domain	EQ-5D	r	r	r	r	r	r	r	r	r
Body Function	Anxiety/Depression	-0.69	-0.48	-0.36	-0.41	-0.33	-0.40	-0.41	-0.54	-0.45
	Pain/Discomfort	-0.37	-0.38	-0.39			-0.50		-0.48	-0.46
Activity	Mobility Self-Care				-0.28	-0.30	-0.40			
					-0.34		-0.34			
Participation	Usual Activities		-0.33	-0.38	-0.32	-0.32	-0.63	-0.12	-0.38	-0.37
SF-36										
Body Function	Mental Health	0.67	0.49	0.36			0.47	0.37	0.54	0.42
	Vitality	0.56	0.44	0.36	0.37	0.34	0.51	0.38	0.50	0.40
	Pain (also part Participation)	0.44	0.47	0.42	0.42	0.40	0.60	0.26	0.53	0.55
Activity	Physical Functioning	0.40	0.42		0.50	0.50	0.59		0.41	0.46
Participation	Social Functioning	0.60	0.50	0.37	0.36	0.33	0.59	0.31	0.56	0.53
	Role Physical	0.48	0.43	0.40	0.43	0.44	0.70	0.18	0.42	0.47
	Role Emotional	0.47	0.37		0.32		0.45	0.16	0.45	0.40

Number of observations between 172 and 182.

r = Pearson correlation.

Correlations are all significant at the .05 level.

Only correlations exceeding .30 are presented excluding those that were hypothesized to be high as marked by the squares.

□ = theoretically assumed overlap.

Table 3 – Pearson correlations between subscales of the EQ-5D and the SF-36.

ICF domain	ICF domain	Body function			Activity	Participation		
	SF-36	Mental Health	Vitality	Pain (also part Participation)	Physical Functioning	Social Functioning	Role Physical	Role Emotional
ICF domain	EQ-5D	r	r	r	r	r	r	r
Body Function	Anxiety/Depression	-0.67	-0.54	-0.42	-0.33	-0.63	-0.42	-0.57
	Pain/Discomfort	-0.39	-0.39	-0.72	-0.43	-0.48	-0.50	-0.44
Activity	Mobility			-0.38	-0.67	-0.30	-0.41	
	Self-Care			-0.45	-0.64	-0.37	-0.30	
Participation	Usual Activities	-0.38	-0.42	-0.55	-0.59	-0.52	-0.68	-0.46

Number of observations between 177 and 179.

r = Pearson's correlation.

Correlations are all significant at the .05 level.

Only correlations exceeding .30 are presented excluding those that were hypothesized to be high as marked by the squares.

□ = theoretically assumed overlap.

Table 4 – Discriminant validity of the BSHS-B subscales.

BSHS-B		No-surgery N = 56	≥1 surgery N = 128	p value	DA hit ratio %
		Mean (SD)	Mean (SD)		
ICF domain Body function	Affect	3.65 (0.69)	3.43 (0.81)	0.07	69.6
	Sexuality	3.82 (0.61)	3.51 (0.82)	0.01*	69.4
	Heat Sensitivity	3.03 (0.92)	2.78 (0.96)	0.10	69.4
Activity	Simple Abilities	3.88 (0.59)	3.69 (0.68)	0.07	69.6
	Hand Function	3.84 (0.61)	3.65 (0.80)	0.07	69.6
Participation	Work	3.67 (0.70)	3.11 (1.15)	0.00*	69.8
	Interpersonal Relationships	3.87 (0.55)	3.77 (0.56)	0.30	69.6
Contextual factors	Body Image	4.59 (1.08)	3.81 (1.45)	0.00*	69.4
	Treatment Regimens	3.68 (0.71)	3.29 (0.96)	0.00*	69.4

N = number of patients, SD = standard deviation, p value = investigation of mean difference between groups *p < 0.05, DA hit ratio = discriminant analysis hit ratio (% correctly classified).

The cross validated classification showed that overall between 69.4 and 69.8% was correctly classified. This is shown in Table 4 as the discriminant analysis (DA) hit ratio.

Regarding the mean EQ-5D subscale all domains showed a statistically significantly higher problem level in the ≥1-surgery group. Both groups presented the most difficulties in the subscales Pain/Discomfort, Anxiety/Depression and Usual Activities, as shown in Table 5. Fig. 1 shows the problem levels of the five dimensions. When tested by Box's M the following subscales showed a significant unequal co-variance between both groups and should be interpreted with caution, i.e., Mobility, Self-Care and Usual Activities. The cross validated classification showed that overall between 69.7 and 70.1% was correctly classified.

The SF-36 subscales were also able to discriminate across severity groups. In the no-surgery group lowest scores were found in the impairments in the subscales Vitality and Mental Health and for the ≥1-surgery group lowest scores were found in subscales Vitality and Mental Health as well as Role Physical

and Role Emotional at 9 months post injury. All subscales detected statistically significant differences between the severity groups. When tested by Box's M the subscales Physical functioning and Social functioning showed a significant unequal co-variance between the groups. The cross validated classification showed that overall between 68.2 and 71.7% was correctly classified.

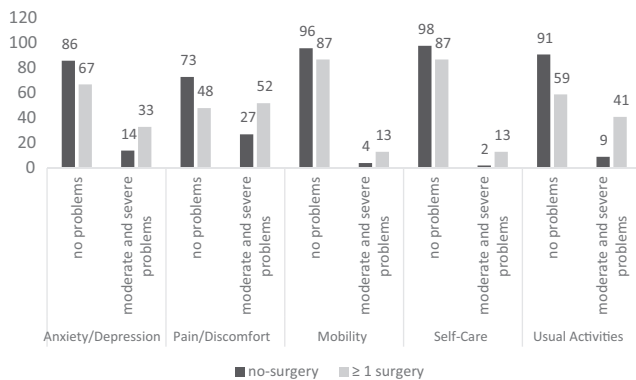
4. Discussion

The present study empirically demonstrates the convergence between the generic EQ-5D and SF-36 and the BSHS-B on subscale level in line with a prior theoretical classification within the ICF framework [4]. The convergence between the two generic scales was stronger than the overlap with the burn specific scale, as reflected by higher correlation coefficients across the subscales. Furthermore, the correlations between the SF-36 and BSHS-B were on average higher as compared to

Table 5 – Discriminant validity of the EQ-5D subscales.

EQ-5D		No-surgery N = 56 Mean (SD)	≥1 surgery N = 129 Mean (SD)	p value	DA hit ratio %
ICF domain					
Body function	Pain/discomfort	0.27 (0.45)	0.55 (0.56)	0.00*	69.7
	Anxiety/depression	0.15 (0.36)	0.36 (0.54)	0.01*	70.1
Activity	Mobility	0.04 (0.19)	0.14 (0.37)	0.05*	70.1
	Self-care	0.02 (0.13)	0.16 (0.41)	0.01*	69.7
Participation	Usual activities	0.09 (0.29)	0.43 (0.54)	0.00*	69.7

N = number of patients, SD = standard deviation, p value = investigation of mean difference between groups *p < 0.05, DA hit ratio = Discriminant analysis hit ratio (% correctly classified).

**Fig. 1 – Distribution of responses in the no-surgery group and ≥1 surgery group to the items of the EQ-5D at 9 months.**

those found in the EQ-5D – BSHS-B comparison. This study also elucidated underlying relationships between domains that provide insight which burn specific domains affect the generic domains of functioning. Furthermore, the results of this study suggest that the generic questionnaires outperform the BSHS-B regarding the ability to differentiate problem levels between burn severity groups. Further study may be required

to establish the nature of these differences, i.e., if these differences are a true reflection of the situation (hand function may be affected in only a subsample of the patients that needed surgery) or whether it is due to the lack of sensitivity of the scale.

This study supports significant content overlap in the generic scales as hypothesized in a prior study using the ICF as the leading framework [4]. Our results are in line with those of Öster et al. [8] in which the EQ-5D index and Visual Analogue Scale (VAS) correlated well with all SF-36 subscales ranging between 0.55 and 0.78. This suggests that similar constructs of HRQOL were measured. Moreover, the correlations between the scales provided insight into underlying problem areas for example, Vitality showed higher correlations with psychological domains such as Anxiety/Depression (EQ-5D) as compared to physical domains suggesting that psychological problems largely affected vitality. This provides valuable insight into underlying problem areas.

The overlap between the EQ-5D and BSHS-B was modest. Only two out of five hypothesized overlapping subscales showed correlations exceeding 0.50, i.e. Anxiety/Depression – Affect and Usual Activities – Work. These results support previous research that showed comparably low associations between BSHS-B subscales and EQ-5D index at 6 and 12 months post burn [8]. Especially Hand Function and

Table 6 – Discriminant validity of the SF-36 subscales.

SF-36		No-surgery N = 57 Mean (SD)	≥1 surgery N = 123 Mean (SD)	p value	DA hit ratio%
ICF domain					
Body Function	Mental health	79.77 (15.81)	70.26 (21.55)	0.00*	68.3
	Vitality	75.20 (17.53)	63.06 (19.83)	0.00*	71.7
	Pain	86.68 (18.16)	75.51 (23.29)	0.00*	68.3
Activity	Physical functioning	94.39 (13.06)	81.95 (21.31)	0.00*	68.3
Participation	Social functioning	92.32 (15.08)	77.34 (25.02)	0.00*	68.3
	Role physical	84.65 (27.85)	59.84 (40.98)	0.00*	68.2
	Role emotional	85.38 (29.56)	67.62 (42.01)	0.01*	68.2

N = number of patients, SD = standard deviation, p value = investigation of mean difference between groups, *p < 0.05, DA hit ratio = Discriminant analysis hit ratio (% correctly classified).

Sexuality respectively showed low correlations with the EQ-5D index and those subscales were relatively independent to other reported health related problems [24]. The findings in this study suggest that both questionnaires measure different components of HRQOL and seem complementary to a large extent. Despite its brevity, the EQ-5D was well able to discriminate between severity groups and further supports its usefulness in burn populations.

Somewhat stronger associations were found between the subscales of the SF-36 and the BSHS-B. This supports the previously demonstrated validity of the SF-36 in burn populations [15]. The BSHS-B subscale Work was strongly associated with several SF-36 subscales in different domains of functioning and showed lower associations with psychological symptoms as was also reported in a prior study [16]. In contrast to earlier findings, however, no evidence of high associations between Heat Sensitivity, Treatment Regimens and Body Image (domain Skin involvement) and role concerns (Social Functioning, Vitality and Mental Health) was detected [16]. The lower associations between appearance related problems and social difficulties in our study may be explained by the lower mean TBSA burned in this group of patients. Contextual factors comprised by the subscales Treatment Regimens and Body Image of the BSHS showed relatively high associations with other ICF domains. This is not that surprising as Contextual Factors are known to influence (facilitate or hinder) all components of functioning [1,25,26].

This study suggests that the generic scales were better able to discriminate problem levels across burn severity as five out of nine BSHS-B subscales (Affect, Heat Sensitivity, Simple Abilities, Hand Function and Interpersonal Relations) did not strongly discriminate between these groups. The finding that affect was more severely disturbed according to the generic scales in the more severely burned group (with subscales Anxiety/Depression in EQ-5D and Mental Health in SF-36) underscores the lower sensitivity in the BSHS-B. These results also support previous findings which reported a superior sensitivity of the SF-36 subscales compared to the BSHS-B over time [15]. However, one could argue that problem levels in some of the scales may have been too small to detect differences across the severity groups. For example Hand function will be affected only in persons with hand burns which will not necessarily be true in all respondents in the more severe burn group. On the other hand, the generic measures may be too broad, also including pre-existing health problems that may have little to do with the burn. Other study methods, e.g., qualitative approaches may be needed to elucidate this issue. Furthermore the short version of the BSHS was used in this study. Longer versions such as the BSHS-Revised (BSHS-R) or the BSHS-Abbreviated (BSHS-A) may score differently on sensitivity.

For use in clinical practice ideally there should be minimal overlap between generic and disease specific instruments to lower the cost of scoring of a number of different measures. However, it is also important to measure all relevant areas of functioning to capture the full impact of a condition. This may add another layer of decision-making and complexity when measuring HRQOL for clinical purposes. This study showed that according to the ICF framework, applying both generic and condition specific measures, did not address the full scope

of problems following burn. None of the generic scales includes Body Structure and other relevant impairment in Body Function (e.g., itch or stigma, coping). Whilst Personal factors can be found in the BSHS-B in the subscale Treatment Regimens and Body Image but this is still limited. Personal factors may include various aspects such as gender, race, age, coping styles, education, profession, past and current experience, individual psychological assets and other health conditions, all of which can affect health and functioning [26]. As illustrated in this study, problems in the participation area were associated with both physical and social problems but the BSHS-B did not make the underlying problems explicit. For example stigmatization may be an underlying problem [27]. These findings indicate there is room for improvement regarding the burn specific instrument both in defining necessary problem areas of functioning currently not included as well as a critical reflection on domains that add little information from a clinical perspective (e.g., treatment regime). In summary, despite the burden to the patient and the clinician it seems necessary to combine a generic measure and a condition specific measure as they are complementary and still incomplete. New technologies such as computerized testing can help limit the burden of duplication in the future [10,28].

Some limitations merit note. This study was part of a larger study design that required patients with adequate cognitive functioning. Older patients and those suffering from cognitive disorders were not included. Proxy assessment of HRQOL in those with mental and cognitive problems would be an alternative for self-reports and should be considered in future studies [2]. In addition persons who were lost to follow-up were statistically younger and may have had a better quality of life. The impact of these biases may limit generalization to the general burn population. The time frame chosen to compare the questionnaires may have influenced the results, as more problems may be experienced at earlier time points. Furthermore, there was a sample size difference between the no surgery and the ≥ 1 surgery group and the measures slightly differ in their timeframe; the EQ-5D examines current health status, the SF-36 uses a 4 week period and the BSHS-B does not specify the timeframe. The sample and timeframe differences may have affected the accuracy of the comparison. Finally, although the correlation boundaries used in this study are accepted in the literature to be classified as high, it can be argued it would be more appropriate to label them moderate. However, as the questionnaires used different time frames and they also differ in their approach (i.e., EQ-5D uses a decision theory approach including single items as indicators of health domains; both SF-36 and BSHS-B use a profile approach that includes multiple items) it was assumed that a generous application of correlation boundaries could be justified.

5. Conclusion

The results of this study illustrate that the ICF framework is useful to classify scales in order to identify overlapping areas. Scales that are largely overlapping unnecessarily increase the burden and should be avoided. The psychometric assessment

of the relationship between the EQ-5D, the SF-36 and the BSHS-B shows that convergence is higher between the generic measures (EQ-5D and SF-36) as compared to the BSHS-B. Both generic scales performed well regarding discriminant validity. Additionally, the ICF framework previously appeared useful to identify gaps regarding domains that are currently overlooked. Consequently, there is room for improvement, particularly regarding the burn-specific measure. This study also illustrated that a short generic scale such as the EQ-5D can provide interesting information but should be combined with a more comprehensive condition-specific scale in order to capture the full impact. Of notice, the use of a coherent timeframe across the scales should be considered to better attune the measurements. Overall, this study supports the view that currently the most frequently used scales to measure functioning fall short to measure the full impact of a burn.

Conflict of interest

All authors declare that they have no conflict of interest to declare.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Acknowledgement

This study was financially supported by the Dutch Burns Foundation, grant numbers 03.09 and 04.2012. This study is the result of collaborative efforts of the psychosocial research group represented by Mr. M. Bremer and Mrs. G. Bakker, Martini Hospital Groningen, the Netherlands, Mrs. A. Boekeelaar, Red Cross Hospital, Beverwijk, The Netherlands, Mrs. H. Hofland and Mrs. A. van de Steenoven, Maasstad Hospital, Rotterdam, the Netherlands, Mrs. L. Braem, ZNA Stuivenberg Antwerp, Belgium, Mrs. A.S. Goemanne, University Hospital Gent, Belgium, Mrs. E. Vandermeulen, Military Hospital, Brussels, Belgium.

REFERENCES

- [1] Van Baar ME, Essink-Bot ML, Oen IMM, Dokter J, Boxma H, Van Beeck EF. Functional outcome after burns: a review. *Burns* 2006;32(1):1–9.
- [2] Van Beeck EF, Larsen CF, Lyons RA, Meerdling WJ, Mulder S, Essink-Bot ML. Guidelines for the conduction of follow-up studies measuring injury-related disability. *J Trauma* 2007;62(2):534–50.
- [3] Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustün B, Stucki G. ICF linking rules: an update based on lessons learned. *J Rehabil Med* 2005;37(4):212–8.
- [4] Meirte J, Van Loey NEE, Maertens K, Moortgat P, Hubens G, Van Daele U. Classification of quality of life subscales within the ICF framework in burn research: identifying overlaps and gaps. *Burns* 2014;40(7):1353–9.
- [5] Brusselaers N, Pirayesh A, Hoeksema H, Verbelen J, Blot S, Monstrey S. Burn scar assessment: a systematic review of objective scar assessment tools. *Burns* 2010;36(8):1157–64.
- [6] Van Loey NEE, Van Beeck EF, Faber BW, Van de Schoot R, Bremer M. Health-related quality of life after burns: a prospective multicentre cohort study with 18 months follow-up. *J Trauma* 2011;72(2):1–8.
- [7] Finlay V, Phillips M, Wood F, Hendrie D, Allison GT, Edgar D. Enhancing the clinical utility of the Burn Specific Health Scale-Brief: not just for major burns. *Burns* 2013;40:328–36.
- [8] Öster C, Willebrand M, Dyster-Aas J, Kildal M, Ekselius L. Validation of the EQ-5D questionnaire in burn injured adults. *Burns* 2009;35(6):723–32.
- [9] Wasiak J, Lee SJ, Paul E, Mahar P, Pfitzer B, Spinks A, et al. Predictors of health status and health-related quality of life 12 months after severe burn. *Burns* 2014;40(4):568–74.
- [10] Ryan CM, Lee AF, Kazis LE, Shapiro GD, Schneider JC, Gorman J, et al. Is real-time feedback of burn-specific patient-reported outcome measures in clinical settings practical and useful? A pilot study implementing the young adult burn outcome questionnaire. *J Burn Care Res* 2015. <http://dx.doi.org/10.1097/BCR.0000000000000287>.
- [11] Mulhern B, Meadows K. The construct validity and responsiveness of the EQ-5D, SF-6D and Diabetes Health Profile-18 in type 2 diabetes. *Health Q Life Outcomes* 2014;12(42):1–10.
- [12] Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value Health* 2007;10(2):S94–105.
- [13] Van Loey NE, Van de Schoot R, Gerdin B, Faber AW, Sjöberg F, Willebrand M. The burn specific health scale-brief: measurement invariant across European countries. *J Trauma Acute Care Surg* 2013;74(5):1321–6. <http://www.euroqol.org>.
- [14] Edgar D, Dawson A, Hankey G, Phillips M, Wood F. Demonstration of the validity of the SF-36 for measurement of the temporal recovery of quality of life outcomes in burns survivors. *Burns* 2010;36(7):1013–20.
- [15] Willebrand M, Kildal M. Burn specific health up to 24 months after the burn—a prospective validation of the simplified model of the burn specific health scale-brief. *J Trauma* 2011;71(1):78–84.
- [16] Brazier J, Walters S, Nicholl J, Kohler B. Using the SF-36 and Euroqol on an elderly population. *Q Life Res* 1996;5:195–204. http://www.rand.org/health/surveys_tools/mos/36-item-short-form/scoring.html.
- [17] Willebrand M, Kildal MA. Simplified domain structure of the burn-specific health scale-brief (BSHS-B): a tool to improve its value in routine clinical work. *J Trauma* 2008;64(6):1581–6.
- [18] Kildal M, Andersson G, Fugl-Meyer AR, Lannerstam K, Gerdin B. Development of a brief version of the Burn Specific Health Scale (BSHS-B). *J Trauma* 2001;51(4):740–6.
- [19] Kim TH, Jo MW, Lee S, Kim SH, Chung SM. Psychometric properties of the EQ-5D-5L in the general population of South Korea. *Q Life Res* 2013;22(8):2245–53.
- [20] Laerd statistics. <https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php>.
- [21] Tyack Z, Simons M, Spinks A, Wasiak J. A systematic review of the quality of burn scar rating scales for clinical and research use. *Burns* 2012;38(1):6–18. <http://dx.doi.org/10.1016/j.burns.2011.09.021>.

-
- [24] Kildal M, Andersson G, Gerdin B. Health status in Swedish burn patients. Assessment utilising three variants of the Burn Specific Health Scale. *Burns* 2002;28(7):639–45.
- [25] <http://www.who.int/classifications/icf/en>.
- [26] Peterson DB. International classification of functioning, disability and health: an introduction for rehabilitation psychologists. *Rehabil Psychogol* 2005;50(2):105–12.
- [27] Lawrence JW, Rosenberg L, Mason S, Fauerbach JA. Comparing parent and child perceptions of stigmatizing behavior experienced by children with burn scars. *Body Image* 2011;8(1):70–3. <http://dx.doi.org/10.1016/j.bodyim.2010.09.004>.
- [28] Ryan CM, Lee AF, Kazis LE, Shapiro GD, Schneider JC, Goverman J, et al. Benchmarks for multidimensional recovery after burn injury in young adults. *J Burn Care Res* 2013;34(3):e121–42. <http://dx.doi.org/10.1097/BCR.0b013e31827e7ecf>.