

# Measuring Patient Needs and Benefits in Dermatology using the Patient Benefit Index 2.0: A Validation Study

Janine TOPP<sup>1</sup>, Matthias AUGUSTIN<sup>1</sup>, Kathrin VON USSLAR<sup>1</sup>, Ramona GOSAU<sup>1</sup>, Kristian REICH<sup>2</sup>, Michael REUSCH<sup>3</sup> and Christine BLOME<sup>1</sup>

<sup>1</sup>German Center for Health Services Research in Dermatology (CVderm), Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), <sup>2</sup>Dermatologikum Berlin and SCIderm Research Institute, and <sup>3</sup>Dermatological Practice Tibarg, Hamburg, Germany

**This study investigated the validity and feasibility of the Patient Benefit Index 2.0 (PBI 2.0), a short instrument to assess patient-relevant treatment benefit. In a cross-sectional study, patients with skin diseases completed the PBI 2.0 alongside instruments on quality of life and disease-specific PBI long versions to assess convergent validity. Feasibility questions appraise comprehensibility, completeness, length, and readability. Data from a longitudinal study were used to explore responsiveness and test-retest reliability. Most patients rated the PBI 2.0 easy to understand, complete, legible, and not too long. The amount of missing values was overall low. In all groups, except for vitiligo, correlation analyses indicated good convergent validity of PBI 2.0. Responsiveness of the PBI 2.0 could not be clearly confirmed. Retest-reliability achieved satisfactory results. Thus, the PBI 2.0 may be a suitable instrument for its use in different skin diseases. Its broad applicability allows for comparisons across diagnosis groups.**

**Key words:** patient-relevant benefit; patient-reported outcomes; validation; skin diseases.

Accepted Oct 11, 2018; Epub ahead of print Oct 11, 2018

Acta Derm Venereol

**Corr:** Janine Topp, German Center for Health Services Research in Dermatology, Institute for Health Services Research in Dermatology and Nursing, University Medical Center Hamburg-Eppendorf, Martinistr. 52, DE-20246 Hamburg, Germany. E-mail: j.topp@uke.de

Patients and healthcare professionals assess patients' health state differently, and assessment results often do not coincide (1–3). Therefore, in order to evaluate treatment benefit comprehensively, the evaluation of the patients' perspective should complement objective measures and outcomes. This demand for patient-relevant benefit assessment is increasingly recognized in clinical encounters as well as in research.

Very commonly, patient-relevant benefit is determined using questionnaires on health-related quality of life (HRQoL), asking patients to rate their impairments due to the disease before and after treatment. A change in HRQoL is interpreted as patient-relevant treatment benefit. This prospective approach, however, is susceptible to recalibration response shift (4), a bias due to changes in the patients' interpretation of the questionnaire and its response scale

## SIGNIFICANCE

Considering the patients' perspective in clinical decision-making is a major goal in healthcare. To assess the patient perspective, valid and reliable instruments on individual needs and treatment benefits are needed. In the dermatological setting, the Patient Benefit Index (PBI) can be used to evaluate patient-relevant treatment benefit. This study analyses and confirms the reliability, feasibility and validity of a short version of the PBI, the PBI 2.0, which is applicable for patients with different skin diseases. The brevity of the instrument may foster its implementation in clinical practice.

(5, 6). This bias is present only in repeated measurements, assessing changes over time. It is not present in retrospective benefit assessment, taking place after treatment only. While potentially being susceptible to recall bias (6), the retrospective assessment has the advantage of measuring benefit close to the actual perception of the patient. This direct rating of benefit by the patient is crucial when assessing treatment satisfaction and estimating treatment adherence (7). Such retrospective approach is taken by the questionnaire Patient Benefit Index (PBI). Furthermore, the PBI has the advantage that, in addition to a benefit assessment, it assesses the importance the individual patient assigns to different treatment benefits and therefore enables the calculation of an importance-weighted global benefit score. The PBI is based on earlier approaches to individualized benefit measurement. One of these approaches is the goal attainment scaling (8), which comprises the formulation of individualized treatment goals with each patient. Another approach is the goal-oriented outcome measurement (9), which includes the determination of individual patient-relevant treatment goals in the field of rehabilitation.

The PBI consists of 2 parts: in the first part, patients rate the perceived importance of different treatment goals listed in the Patient Needs Questionnaire (PNQ). In the second part, the Patient Benefit Questionnaire (PBQ), patients rate the achievement of these treatment goals. Usually, the PNQ is completed before starting a new treatment, and the PBQ at a time when an effect of the treatment is expected and treatment benefit will be assessed (10).

Thus far, a range of different PBI instruments have been developed and validated for various skin diseases. A short version of the PBI, the PBI 2.0, has been developed based

on 9 disease-specific PBI versions to address needs and benefits of patients with superordinate skin diseases (11). PBI versions of the following patient groups and indications have been used for the development of the short version: chronic inflammatory skin diseases (psoriasis, acne, atopic dermatitis) (12, 13), vitiligo (14), rosacea, venous diseases, chronic pruritus (15), chronic hand eczema (16), chronic wounds (17), chronic wounds under ultrasound treatment (18), and cosmetic indications (19). After categorizing all items of these different PBI versions, one or more items that summarize the content of each category were drafted for its use in the PBI 2.0. The new items had to be applicable for skin diseases in general and had to cover the content of diagnosis-specific items on a superordinate level. This was done by using more generic expressions, such as “skin problems” instead of its specifications, e.g. “itching” or “skin reddening”. The appropriateness of the PBI 2.0 items was evaluated in 16 semi-structured interviews with patients having a chronic skin disease. Items were revised subsequently based on the patient evaluation, resulting in a final version of the PBI 2.0 (11).

The PBI 2.0 has the advantage of being only approximately half as long as the diagnosis-specific versions, thus decreasing administrative burden for the patient, and of being applicable to patients with different skin diseases, thereby allowing for comparison across diagnosis groups.

The aim of this study was to evaluate the validity and feasibility of the German version of the PBI 2.0 in patients with different chronic skin diseases.

## METHODS

### *Patient Benefit Index 2.0*

The PBI 2.0 comprises 12 items describing treatment goals of patients with skin diseases. In a first step, these treatment goals are rated on a 5-point Likert-scale (0=“not at all”, 4=“very much”) according to their perceived relevance (PNQ). Alternatively, patients can state that a goal “does not apply” to them. In a second step, patients are asked to express the achievement of these 12 treatment goals (treatment benefit) on the same 5-point Likert-scale (PBQ), again with the opportunity to state that the goal did not apply. The PNQ and PBQ responses are converted to a weighted PBI global score, ranging from 0=“no benefit” to 4=“maximum benefit”. According to the manual, the PBI global score can be calculated if a patient provided valid data on the PNQ (importance rating) and the PBQ (benefit rating) for at least 75% of the respective treatment goals.

### *Data sources for the validation*

Two validation studies were conducted: a cross-sectional and a longitudinal study. The items of the German version of the PBI 2.0 have been translated to English for this manuscript, but the translation has not yet been formally validated.

**Cross-sectional validation study.** A convenience sample of German-speaking adults with chronic skin diseases was recruited at the dermatological outpatient clinic of the German Center for Health Services Research in Dermatology (CVderm), University Medical Center Hamburg-Eppendorf. Patients with vitiligo were contacted via an email newsletter of the German Vitiligo Associa-

tion (*Deutscher Vitiligo Verein e.V.*). A total of 814 questionnaires were dispensed, including 360 questionnaires for patients with psoriasis, 120 for patients with leg ulcer, 272 for patients with atopic dermatitis, and 62 for patients with vitiligo. The questionnaire set contained the PBI 2.0 and the following instruments:

- The *Dermatology Life Quality Index (DLQI)*: this validated and frequently used measure consists of 10 items assessing different aspects of skin disease-specific HRQoL. The items can be summed up to a total score ranging from 0 to 30, where higher scores indicate lower quality of life (20, 21).
- The *EQ-5D-3L*: this questionnaire on subjective health is a widely used preference-based measure, assessing 5 dimensions of health. Patients state their perceived problems in the respective dimension on 3 possible levels, resulting in 243 possible health states. The health state can then be converted into a single summary index by attaching weights to the levels in each dimension (22). Validity and responsiveness of the EQ-5D-3L in skin diseases has been shown to be good overall (23).
- The respective *diagnosis-specific versions of the PBI*: we used the versions PBI-S in atopic dermatitis and psoriasis (12), PBI-Vit in vitiligo (14), and PBI-W in leg ulcers (17). All have a similar structure, as described above, and have been evaluated for their validity and feasibility (10, 12, 14, 17).

**Longitudinal validation study.** In the longitudinal study, data from patients with chronic skin diseases were collected at 3 time points: before initiation of a new treatment (t0), approximately 1 week (t1), and 3 months after treatment onset (t2). A new treatment included the onset of a completely new therapy as well as changes in existing treatment procedures, such as a modification in the dosage of a medication. Data collection took place at the CVderm, at 3 other divisions of the University Medical Center Hamburg-Eppendorf, and at the venous outpatient clinic of the Tabea hospital, all located in Hamburg, Germany. To assess the applicability of the PBI 2.0 in a diverse population, patients from 9 different diagnosis groups were recruited: psoriasis, venous diseases, chronic pruritus, leg ulcer, other chronic wounds, atopic dermatitis, arterial diseases, chronic hand eczema, and actinic keratosis. We aimed for recruiting 20 patients in each group and analysed them in a combined sample, not differentiating between diagnosis groups.

The t0 and t2 questionnaires contained the PBI 2.0 and the following instruments:

- The *DLQI* (described above).
- The *EQ-5D-5L*: This is a newer version of the EQ-5D-3L, which was described above, using a 5-level instead of a 3-level response scale (24).
- The *Freiburg Life Quality Assessment (FLQA)*: This validated measure assesses dermatology-specific HRQoL and consists of 28 items assigned to 6 scales: physical complaints, everyday life, social life, emotional status, treatment and satisfaction (25).

The t2 questionnaire contained additional single-item questions on the overall HRQoL and the benefit of the received treatment, which could be answered on 5-point Likert-scales. At t1, patients were asked to complete the PBI 2.0 only.

### *Statistical analyses*

Distributional characteristics (mean and standard deviation) were determined for the PBI 2.0 global scores separately for both data sources. The amount of missing values was analysed and served as an indicator for the acceptance of the measure.

In order to test for convergent validity, we computed Pearson correlations. Cross-sectional data were used to assess convergent validity of the PBI 2.0 global score with the respective global scores of the DLQI and the EQ-5D-3L. In addition, the correlation of the PBI 2.0 with the respective disease-specific PBI versions was determined.

For the analysis of responsiveness, we defined partial correlations of the PBI 2.0, calculated from the PNQ at t0 and the PBQ at t2, and the global scores of the DLQI, the EQ-5D-5L, and the FLQA at t2 controlling for the respective global scores at t0. Furthermore, Pearson correlations of the PBI 2.0 with the non-standardized questions on retrospective patient-relevant treatment benefit were computed.

We did not consider internal consistency, as the PBI 2.0 is based on a formative model, implying that items together form the construct, not expecting single items to be highly correlated (26, 27).

Test-retest reliability of each item of the PNQ was assessed based on the longitudinal data collected at t0 and 1 week later at t1. The reason for only assessing the test-retest reliability of the PNQ was that a new treatment was initiated at t0. Therefore, benefit assessment (PBQ) at t0 referred to the treatment patients received before t0, whereas benefit assessment at t1, one week after the onset of the new therapy, referred to the new treatment. The divergent reference frame did not allow for an evaluation of retest-reliability for the second part of the questionnaire (PBQ). The intraclass correlation coefficient (ICC) and Cohen's  $\kappa$  were calculated for each of the 12 items of the PNQ. The ICC is based on the assumption of metric data, which is why we allocated the response option "does not apply to me" to "not at all relevant". Cohen's  $\kappa$  considers agreement of response options, which is why we determined the response option "does not apply to me" as a separate response option. Test-retest reliability was assessed if the period between t0 and t1 did not exceed 7 days, to ensure a limited time frame within which the construct was not expected to have changed considerably.

Feasibility questions from the cross-sectional study were analysed descriptively, and free-text comments were grouped and evaluated by content.

Data analyses were conducted with IBM SPSS Statistics, Version 23.

### Ethics

The study was carried out in accordance with the Declaration of Helsinki. All patients provided informed consent.

## RESULTS

### Cross-sectional validation study

**Patient characteristics.** Overall, 379 questionnaires were completed and returned (response rate 46.6%) including

153 from patients with psoriasis, 116 from patients with leg ulcers, 64 from patients with atopic dermatitis, and 46 from patients with vitiligo.

In patients with atopic dermatitis and leg ulcers, male patients were predominant with 53.1% and 53.4%, respectively, whereas in psoriasis and vitiligo there were more female patients with 54.2% and 65.2% (Table I). While the school educational level of most patients with atopic dermatitis and vitiligo was high, patients with leg ulcers and psoriasis more often had a middle educational level. Patients with leg ulcers had a mean age of 65.9 years, while the other 3 groups were younger, with a mean age of 41.5–50.0 years. Mean disease duration, defined as years since first diagnosis, was highest in patients with atopic dermatitis (26.3 years) and lowest in patients with leg ulcers (6.5 years) (Table I).

**Distribution of PBI 2.0 and missing values.** Mean patient-relevant benefit according to the PBI 2.0 global score was highest in patients with psoriasis ( $3.0 \pm 1.0$ ,  $n = 136$ ), followed by patients with leg ulcers ( $2.7 \pm 1.2$ ,  $n = 102$ ) and atopic dermatitis ( $2.3 \pm 0.8$ ,  $n = 59$ ). The lowest treatment benefit was reported by patients with vitiligo ( $1.5 \pm 1.3$ ,  $n = 15$ ) (Table I). In this group, the majority of patients did not receive any treatment, which made benefit assessment impossible (PBQ) and led to a sample size of 15 patients only.

All but 2 patients completed the PNQ on individual patient needs, at least partially. The PBQ on treatment benefits, was not completed by 65 of 379 patients (16.9%); in 64 of them this was because they currently did not receive any treatment, making benefit evaluation impossible (Table II).

In those patients who at least partially completed the PBI 2.0, the rate of missing values ranged from 0.0% to 6.7%, depending on item and patient group. The highest rate was found for the PBQ item number 2 "having a treatment with few side-effects" (Table II). The mean number of missing values per patient in the PNQ ranged

**Table I. Sociodemographic and clinical data of cross-sectional and longitudinal validation studies**

	Cross-sectional validation study				Longitudinal validation study $n = 199$
	Psoriasis $n = 153$	Leg ulcers $n = 116$	Atopic dermatitis $n = 64$	Vitiligo $n = 46$	
Sex, %					
Female	54.2	44	43.8	65.2	51.8
Male	43.8	53.4	53.1	34.8	47.2
Not specified	2	2.6	3.1	0	1.0
Age, years, mean $\pm$ SD	50 $\pm$ 13.8	65.9 $\pm$ 14.7	41.5 $\pm$ 13.8	46.7 $\pm$ 11.7	55.3 $\pm$ 18.3
Median (range)	51 (17–77)	70 (15–93)	42 (22–75)	46 (26–80)	58 (19–92)
Not specified, $n$	4	6	1	–	6
Educational level (%)					
Low (primary education)	28.8	40.5	15.6	8.7	27.1
Middle (secondary education)	34.0	31.9	28.1	26.1	34.7
High (high school degree)	32.0	23.3	51.6	63.0	37.2
Not specified	5.2	4.3	4.7	2.2	1.0
Years since first diagnosis, mean $\pm$ SD	22.0 $\pm$ 13.8	6.5 $\pm$ 8.5	26.3 $\pm$ 15.0	17.8 $\pm$ 11.7	11.1 $\pm$ 13.9
Median (range)	20 (1–58)	3 (1–41)	26.5 (1–71)	16 (3–48)	5 (0–69)
Not specified, $n$	4	14	2	1	20
PBI 2.0 global score, mean $\pm$ SD	3.0 $\pm$ 1.0	2.7 $\pm$ 1.2	2.3 $\pm$ 0.8	1.5 $\pm$ 1.3	2.5 $\pm$ 1.2

SD: standard deviation; PBI 2.0: Patient Benefit Index 2.0.

**Table II. Missing values in the Patient Benefit Index 2.0 (PBI 2.0) in the cross-sectional validation study**

	Psoriasis (n = 153)		Leg ulcers (n = 116)		Atopic dermatitis (n = 64)		Vitiligo (n = 46)	
	PNQ	PBQ	PNQ	PBQ	PNQ	PBQ	PNQ	PBQ
PBI 2.0 completed, at least partially, n	153	136	114	104	64	59	46	15
PBI 2.0 not completed, n (%)	–	17 (11.1)	2 (1.7)	12 (10.3)	–	5 (7.8)	–	31 (67.4)
Missing values per item, %								
1. to have less physical discomfort (e.g. pain, burning, itching)	0	0	0	2.9	0	0	0	0
2. to have fewer side-effects	0.7	0.7	1.8	1.9	0	1.7	0	6.7
3. to be able to sleep better	1.3	0.7	0	2.9	1.6	0	0	0
4. to be happier and more satisfied	0.7	1.5	0	1	0	0	0	0
5. to feel more comfortable in contact with others	0.7	1.5	4.4	1.9	1.6	0	0	0
6. to feel more attractive	0.7	0.7	4.4	1	0	0	0	0
7. to be able to engage in normal leisure activities	0.7	2.2	1.8	5.8	0	0	0	0
8. to have a normal everyday life (work, housekeeping, daily activities)	0.7	0.7	0	0	0	1.7	0	0
9. to be less of a burden to relatives and friends	2	0.7	0.9	1	0	3.4	0	0
10. to spend less money on treatment	1.3	1.5	1.8	1.9	0	1.7	0	0
11. to have less bother with treatment	0.7	1.5	0	1	0	0	0	0
12. to better foresee and control the disease	1.3	2.9	1.8	2.9	0	0	0	0
Mean number of missing values per patient	0.1	0.15	0.17	0.24	0.03	0.08	0	0.07

PNQ: Patient Needs Questionnaire; PBQ: Patient Benefit Questionnaire.

from 0.0 items for patients with vitiligo to 0.17 items for patients with leg ulcer. In the PBQ, the mean number of missing values ranged from 0.07 items for patients with vitiligo to 0.24 items for patients with leg ulcer (Table II).

#### Longitudinal validation study

**Patient characteristics.** The baseline data set (t0) contained 199 patients. One hundred and seventy-six of them (88.4%) returned the first follow-up questionnaire (t1) and 150 (75.4%) the second (t2). Patients could be allocated to the following groups (t0): psoriasis (n = 30), venous diseases (n = 25), chronic pruritus (n = 24), leg ulcer (n = 22), other chronic wounds (n = 23), atopic dermatitis (n = 22), arterial diseases (n = 21), chronic hand eczema (n = 20), and actinic keratosis (n = 12). Thus, the intended sample size of 20 patients per group was achieved for all but the patient group of actinic keratosis. Of all patients 52.3% were female and mean age was 55.3 years. On average, first diagnoses had been made 11.1 years previously (Table I).

Disease-specific HRQoL of all patients determined with the DLQI was relatively good at t0 ( $7.9 \pm 6.0$ ; possible range: 0–30 where higher scores indicate lower HRQoL) and seemed slightly better at t2 ( $7.6 \pm 6.1$ ). Slight improvements were also observed for the EQ-5D-5L (t0:  $71.8 \pm 24.3$ ; t2:  $73.6 \pm 23.6$ ), while the FLQA score remained the same (t0:  $2.4 \pm 0.5$ , t2:  $2.4 \pm 0.5$ ).

**Distribution of PBI 2.0 and missing values.** According to the PBI 2.0 global score, mean patient-relevant benefit of the treatment, that was initiated at t0, was  $2.5 \pm 1.2$  (n = 134). This calculation was based on the rating of individual needs (PNQ) at the onset of the new treatment at t0 and evaluation of its benefit (PBQ) 3 months after at t2.

Missing value analyses were carried out separately for the PNQ and the PBQ at the 3 time points. The PNQ was completed, at least partially, by 98–100% of the sample, depending on the time point. Non-completion of the PBQ was more prevalent. At t0, 30.2% of all patients did not

answer the PBQ. This high amount of non-completion at t0 is presumably because a new treatment was introduced at t0 only, not allowing for direct benefit assessment. At t1 the rate of non-completion of the PBQ decreased to 13.6%. At this time point treatment had been started only one week previously, which may explain why patients already felt incapable of evaluating benefit. At t2, 6.7% of patients did not complete the PBQ.

In those patients who at least partially completed the questionnaire, the rate of missing values per item ranged from 0.0% to 5.9%. The mean number of missing values per patient ranged from 0.2 to 0.3 items, depending on the time point and part of the PBI 2.0 (Table III).

#### Convergent validity

The correlation of the PBI 2.0 with the skin-specific DLQI ranged from  $r=0.29$  (leg ulcers) to  $r=0.58$  (atopic dermatitis). Correlations with the generic EQ-5D tended to be lower with  $r=0.20$  (psoriasis) to  $r=0.54$  (atopic dermatitis). Most correlations were highly significant ( $p < 0.01$ ), except for the subgroup of patients with vitiligo. Here, a PBI 2.0 global score could be calculated for only 15 patients because of the high number of patients who did not complete the PBQ due to a lack of current treatment (Table IV).

High correlations were found between PBI 2.0 global scores and global scores of the respective disease-specific PBI long versions, ranging from  $r=0.69$  in patients with leg ulcers to  $r=0.84$  in patients with atopic dermatitis. Convergent validity of PBI 2.0 and disease-specific PBI versions was mostly very similar, with a maximum difference of 0.05 points in correlation. In patients with leg ulcers, convergent validity regarding the DLQI was 0.17 points higher in the PBI 2.0 compared with the PBI-W. Only in patients with vitiligo, convergent validity of the PBI 2.0 regarding both criteria was 0.12 points lower than in the vitiligo-specific PBI-Vit. All results are shown in Table IV.

**Table III. Missing values in the Patient Benefit Index 2.0 (PBI 2.0) in the longitudinal validation study**

	t0 (n = 199)		t1 (n = 176)		t2 (n = 150)	
	PNQ	PBQ	PNQ	PBQ	PNQ	PBQ
PBI 2.0 completed, at least partially, n	198	139	176	152	147	140
PBI 2.0 not completed, n (%)	1 (0.5)	60 (30.2)	-	24 (13.6)	3 (2.0)	10 (6.7)
Missing values per item, %						
1. to have less physical discomfort (e.g. pain, burning, itching)	1.0	0.7	0	0.7	0.7	1.4
2. to have fewer side-effects	1.5	3.6	1.7	5.9	2.0	2.1
3. to be able to sleep better	2.5	2.2	3.4	2.0	2.7	3.6
4. to be happier and more satisfied	5.1	5.0	2.3	3.9	2.0	3.6
5. to feel more comfortable in contact with others	4.0	2.9	1.7	2.6	2.0	2.1
6. to feel more attractive	2.0	2.9	2.3	2.6	3.4	2.9
7. to be able to engage in normal leisure activities	3.5	5.0	1.7	2.6	2.0	2.1
8. to have a normal everyday life (work, housekeeping, daily activities)	2.5	1.4	0.6	0	1.4	1.4
9. to be less of a burden to relatives and friends	3.0	3.6	2.3	0.7	3.4	1.4
10. to spend less money on treatment	2.0	2.2	1.7	2.6	0.7	2.9
11. to have less bother with treatment	1.5	1.4	0.6	0.7	2.0	2.1
12. to better foresee and control the disease	2.0	1.4	5.7	0.7	1.4	0.7
Mean number of missing values per patient	0.3	0.3	0.2	0.3	0.2	0.3

PNQ: Patient Needs Questionnaire; PBQ: Patient Benefit Questionnaire; t0: before initiation of a new treatment; t1: approximately 1 week after treatment onset; t2: 3 months after treatment onset.

### Responsiveness

Responsiveness of the PBI 2.0 considering changes in HRQoL over time could not explicitly be confirmed. While a significant correlation of the PBI 2.0 (PNQ at t0, PBQ at t2) and change in DLQI from t0 to t2 was found ( $r=-0.36$ ;  $p<0.001$ ), correlations with changes in the EQ-5D-5L and the FLQA were not statistically significant. This may be explained by the low variance of both measures between t0 and t2, i.e. the low degree of changes. The EQ-5D-5L global score, on average, increased by 1.8 points (range: 0–100) while the FLQA changed by 0.01 points (range 1–5).

In contrast, correlations of the PBI 2.0 with 6 questions on change in HRQoL and patient-relevant treatment benefit were statistically significant ( $r=0.46$  to  $0.57$ ) (Table V).

**Table IV. Convergent validity of the Patient Benefit Index 2.0 (PBI 2.0) based on cross-sectional validation study**

	Psoriasis	Leg ulcers	Atopic dermatitis	Vitiligo
Correlation of PBI 2.0 with DLQI <sup>a</sup>				
r	-0.48	-0.29	-0.58	-0.49
p	<0.001	0.005	<0.001	0.062
n	133	92	57	15
Correlation of PBI 2.0 with EQ-5D-3L				
r	0.20	0.36	0.54	0.19
p	0.026	<0.001	<0.001	0.508
n	130	96	58	15
Correlation of PBI 2.0 with disease-specific PBI <sup>b</sup>				
r	0.83	0.69	0.84	0.93
p	<0.001	<0.001	<0.001	<0.001
n	125	94	53	13
Correlation of disease-specific PBI with DLQI <sup>a</sup>				
r	-0.52	-0.12	-0.59	-0.61
p	<0.001	0.243	<0.001	0.009
n	125	91	53	17
Correlation of disease-specific PBI with EQ-5D-3L				
r	0.16	0.32	0.49	0.31
p	0.071	0.002	<0.001	0.225
n	122	93	53	17

<sup>a</sup>Negative correlations result from the fact that high values in the Dermatology Life Quality Index (DLQI) represent low quality of life, whereas high values in the PBI 2.0 represent high benefit; <sup>b</sup>psoriasis and atopic dermatitis=PBI-S; leg ulcers=PBI-W; vitiligo=PBI-Vit. r: Pearson correlation coefficient; EQ-5D-3L: European Quality of Life Questionnaire 5 dimensions 3 levels.

### Test-retest reliability

Test-retest reliability of the PNQ was satisfactory. Cohen's  $\kappa$  values between 0.41 and 0.55 indicated sufficient agreement between response options. Allocating the response option "does not apply to me" to the category "not at all relevant" and interpreting scales of the items as metric, resulted in ICC from 0.58 to 0.79, which suggests fair to good overall test-retest reliability of the PNQ (Table VI) (26, 28).

### Feasibility

In the feasibility questionnaire of the cross-sectional validation study, 96.6% of patients stated that the instructions

**Table V. Responsiveness of the Patient Benefit Index 2.0 (PBI 2.0) global score based on longitudinal validation study**

	r	p-value	n
Correlation of PBI 2.0 with $\Delta$ DLQI <sup>a</sup>	-0.36	<0.001	122
Correlation of PBI 2.0 with $\Delta$ EQ-5D-5L	0.15	0.10	117
Correlation of PBI 2.0 with $\Delta$ EQ VAS	0.20	0.03	127
Correlation of PBI 2.0 with $\Delta$ FLQA	-0.12	0.20	120
Subscale: physical complaints	-0.14	0.11	125
Subscale: everyday life	-0.11	0.22	126
Subscale: social life	-0.13	0.14	125
Subscale: emotional status	-0.15	0.09	123
Subscale: treatment	-0.19	0.06	96
Subscale: satisfaction	0.34	<0.001	123
Correlation of PBI 2.0 with non-standardized items on retrospective treatment benefit			
Physical condition	0.56	<0.001	132
Mental condition	0.55	<0.001	131
Everyday life	0.50	<0.001	129
Social interaction	0.47	<0.001	127
Leisure time activities	0.49	<0.001	129
Overall quality of life	0.48	<0.001	129
Overall effectiveness of the treatment	0.53	<0.001	129
Recommendation of the treatment	0.46	<0.001	129

Global scores (mean (standard deviation)) DLQI: t0=7.93 (6.04), t2=7.61 (6.08); EQ-5D-5L: t0=71.79 (24.27), t2=73.60 (23.58); FLQA: t0=2.41 (0.47), t2=2.42 (0.48).

<sup>a</sup>Negative correlations result from the fact that high values in the DLQI represent low quality of life, whereas high values in the PBI 2.0 represent high benefit. r: Pearson correlation coefficient; p: level of significance; DLQI: Dermatology Life Quality Index; EQ-5D-5L: European Quality of Life Questionnaire 5 dimensions 5 levels; EQ VAS: European Quality of Life Questionnaire - visual analogue scale; FLQA: Freiburg Life Quality Assessment.

**Table VI. Retest-reliability of the Patient Needs Questionnaire (PNQ) based on longitudinal validation study**

Item	$\kappa$	$p$ -value	ICC	$p$ -value	$n$
1. to have less physical discomfort (e.g. pain, burning, itching)	0.50	<0.001	0.59	<0.001	133
2. to have fewer side-effects	0.45	<0.001	0.58	<0.001	132
3. to be able to sleep better	0.55	<0.001	0.81	<0.001	131
4. to be happier and more satisfied	0.46	<0.001	0.61	<0.001	128
5. to feel more comfortable in contact with others	0.48	<0.001	0.62	<0.001	131
6. to feel more attractive	0.50	<0.001	0.71	<0.001	131
7. to be able to engage in normal leisure activities	0.47	<0.001	0.64	<0.001	130
8. to have a normal everyday life	0.41	<0.001	0.74	<0.001	131
9. to be less of a burden to relatives and friends	0.45	<0.001	0.69	<0.001	127
10. to spend less money on treatment	0.51	<0.001	0.79	<0.001	132
11. to have less bother with treatment	0.41	<0.001	0.64	<0.001	133
12. to better foresee and control the disease	0.48	<0.001	0.67	<0.001	133

$\kappa$ : Cohen's Kappa;  $p$ : level of significance; ICC: intra-class correlation coefficient.

within the PBI 2.0 were easy to understand (1.6% not easy, 1.8% missing), same as the single items (94.7% yes, 2.9% no, 2.4% missing). The PBI 2.0 was perceived as not too long by 96.3% of patients (1.8% too long, 1.8% missing). 97.4% stated that the questionnaire was legible (1.6% not legible, 1.1% missing).

Patients were asked if they had difficulties deciding on an answer to particular items of the PBI 2.0; 13.5% responded with "yes", 84.4% with "no" (2.1% missing). The percentage of patients who had reported difficulties was highest in the vitiligo subgroup (17.4%).

The question regarding if any important treatment goal was missing in the PBI 2.0 was answered "yes" by 14.8% of the overall sample (82.6% no, 2.6% missing). Approval of this question was 10.5% in psoriasis, 12.9% in leg ulcers and 14.1% in atopic dermatitis; in patients with vitiligo, it was markedly higher, at 34.8%. According to their free text responses, patients with vitiligo particularly regarded treatment goals on attractiveness as insufficiently covered by the PBI 2.0 (less visibility of vitiligo areas; feeling attractive even when lightly dressed). Further goals named as missing related to psychological burden (mental problems; "the psychological component") and the possibility of exposure to UV. Some statements on missing goals related to the acceptance of vitiligo as a disease by society and by health insurance funds. Achievement of these goals, however, is not part of the construct of benefit as measured by the PBI, which intends to quantify patient-relevant benefit of single treatments, but not issues with care or society as a whole.

## DISCUSSION

This study investigated the feasibility and validity of the PBI 2.0. The intention was to evaluate an abbreviated, generic version of the PBI that is applicable for a wider range of patients with skin diseases and therefore allows for comparability across diagnosis groups. The results of this study demonstrate good convergent validity of the PBI

2.0. Test-retest reliability was satisfactory, while responsiveness of the PBI 2.0 could only partly be confirmed. The majority of patients rated the PBI 2.0 as comprehensible, feasible, and complete by, a finding supported by a low rate of missing values in the questionnaire.

The overall positive results considering convergent validity, general feasibility and completeness could not be clearly confirmed for patients with vitiligo. In particular, important treatment goals relating to attractiveness and psychological burden may not be sufficiently covered by the PBI 2.0. Furthermore, for patients with vitiligo the PBI 2.0 was not considered valid regarding the criterion of HRQoL as the respective disease-specific PBI, the PBI-Vit. As conclusions on convergent validity of the PBI 2.0 for this patient group are drawn based on a total sample of 15 patients, the results need to be interpreted with caution.

While convergent validity of the PBI 2.0 with respective disease-specific PBI versions could be confirmed, convergent validity was noticeable lower when comparing the PBI 2.0 with instruments on general health (EQ-5D-5L) and disease-specific HRQoL (DLQI). As comparator measures, we chose measures beyond the symptom level targeting everyday life. We expected lower correlations with EQ-5D-5L and DLQI as their constructs of general health and HRQoL differ from the underlying construct of the PBI (10, 21, 24). To date, no comparator measure other than the respective long PBI version is known targeting the same construct as the PBI 2.0 (treatment needs and benefits). Despite this, our analyses confirmed that the PBI 2.0 was as valid regarding the criterion of HRQoL as the respective disease-specific PBI versions (12, 14, 17), which are approximately twice as long.

As for the analyses of convergent validity, analyses on responsiveness of the PBI 2.0 are limited. A variance of the comparator measures over time would have been needed to allow for interpretations on responsiveness (29, 30). We could only find low variance from t0 to t2 for all 3 comparator measures; DLQI, EQ-5D-5L and FLQA. Significant correlations could be identified for the PBI 2.0 with change in the DLQI only. A reason for the low variance could be that patients already met the inclusion criteria of the study if minor changes in the treatment were made. A change in, for example, the dosage of a medication or the type of wound dressing was sufficient for participation. In these cases no major changes in the HRQoL were expected. Therefore, further research is needed to address responsiveness of the PBI 2.0.

Knowing about the limitations regarding interpretation of validity parameters, it is worth mentioning some strengths of the study. Our validation is based on data of 2 fairly large samples: the cross-sectional study includes data of 379 patients, the longitudinal study of 199 patients. Moreover, the samples include heterogeneous groups of patients with various dermatological conditions. This supports the applicability of the PBI 2.0 for superordinate skin diseases, which was the initial purpose of the

developed short form of the PBI. Furthermore, we aimed to develop a short instrument for which applicability in clinical practice is less of a hurdle. After confirming the overall validity of the PBI 2.0, we encourage researchers and healthcare professionals to use it for standardized benefit assessment. Further research is needed to address the responsiveness and test–retest reliability of the PBQ on treatment benefit.

### Conclusion

The German version of the PBI 2.0, a questionnaire on patient-relevant benefit in different skin diseases that is only half as long as the respective diagnosis-specific PBI versions, proved to be valid and feasible for its use in skin diseases. An exception is the target group of vitiligo, for which the instrument may not cover all relevant treatment goals.

### ACKNOWLEDGEMENTS

This study was supported by the German Federal Ministry of Education and Research (BMBF) within the context of the Hamburg Center for Health Economics (HCHE), grant number 01EH1101.

MA is licence holder of the Patient Benefit Index. All other authors do not have any conflicts of interest.

### REFERENCES

- Rothwell PM, McDowell Z, Wong CK, Dorman PJ. Doctors and patients don't agree: cross sectional study of patients' and doctors' perceptions and assessments of disability in multiple sclerosis. *BMJ* 314: 1580–1583.
- Robinson JA, Crawford GB. Do palliative patients and carers agree about patients' psychological functioning? *Palliat Support Care* 2010; 8: 69–74.
- Stokes JW, Wanderer JP, McEvoy MD. Significant discrepancies exist between clinician assessment and patient self-assessment of functional capacity by validated scoring tools during preoperative evaluation. *Perioper Med* 2016; 5: 18.
- Schwartz CE, Bode R, Repucci N, Becker J, Sprangers MAG, Fayers PM. The clinical significance of adaptation to changing health: a meta-analysis of response shift. *Qual Life Res* 2006; 15: 1533–1550.
- Uebel PA, Peeters Y, Smith D. Abandoning the language of "response shift": a plea for conceptual clarity in distinguishing scale recalibration from true changes in quality of life. *Qual Life Res* 2010; 19: 465–471.
- Blome C, Augustin M. Measuring change in quality of life: bias in prospective and retrospective evaluation. *Value Health* 2015; 18: 110–115.
- Augustin M, Holland B, Dartsch D, Langenbruch A, Radtke MA. Adherence in the treatment of psoriasis: a systematic review. *Dermatology* 2011; 222: 363–374.
- Kiresuk TJ, Sherman RE. Goal attainment scaling: a general method for evaluating comprehensive community mental health programs. *Community Ment Health J* 1968; 4: 443–453.
- Gerdes N. Rehabilitationseffekte bei "zielorientierter Ergebnismessung". *Ergebnisse der IRES-ZOE-Studie 1996/97. Deutsche Rentenversicherung* 1998; 3–4: 217–237.
- Augustin M, Radke MA, Zschocke I, Blome C, Behechtnejad J, Schäfer I, et al. The patient benefit index: a novel approach in patient-defined outcomes measurement for skin diseases. *Arch Dermatol Res* 2009; 301: 561–571.
- Blome C, Usslar K von, Augustin M. Feasibility of using qualitative interviews to explore patients' treatment goals: experience from dermatology. *Patient* 2016; 9: 261–269.
- Feuerhahn J, Blome C, Radke MA, Augustin M. Validation of the patient benefit index for the assessment of patient-relevant benefit in the treatment of psoriasis. *Arch Dermatol Res* 2012; 304: 433–441.
- Augustin M, Reich C, Schäfer I, Zschocke I, Rustenbach SJ. Entwicklung und Validierung eines neuen Instrumentes zur Erfassung patientendefinierten Nutzens in der Therapie der Akne. *J Dtsch Dermatol Ges* 2008; 6: 1–8.
- Augustin M, Gajur AI, Reich C, Rustenbach SJ, Schäfer I. Benefit evaluation in vitiligo treatment: development and validation of a patient-defined outcome questionnaire. *Dermatology* 2008; 217: 101–106.
- Blome C, Augustin M, Siepmann D, Phan NQ, Rustenbach SJ, Ständer S. Measuring patient-relevant benefits in pruritus treatment: development and validation of a specific outcomes tool. *Br J Dermatol* 2009; 161: 1143–1148.
- Blome C, Maares J, Diepgen T, Rustenbach SJ, Augustin M. Measurement of patient-relevant benefits in the treatment of chronic hand eczema – a novel approach. *Contact Dermatitis* 2009; 61: 39–45.
- Augustin M, Blome C, Zschocke I, Schäfer I, Koenig S, Rustenbach SJ. Benefit evaluation in the therapy of chronic wounds from the patients' perspective – development and validation of a new method. *Wound Repair Regen* 2012; 20: 8–14.
- Herberger K, Franzke N, Blome C, Kirsten N, Augustin M. Efficacy, tolerability and patient benefit of ultrasound-assisted wound treatment versus surgical debridement: a randomized clinical study. *Dermatology* 2011; 222: 244–249.
- Augustin M, Schäfer I, Rabini S, Lee-Seifert C, Radke MA, Rustenbach SJ. Assessing the value of supportive skin care: development and validation of an instrument for evaluating patient-relevant benefit. *Dermatology* 2009; 218: 255–259.
- Basra MK, Fenech R, Gatt RM, Salek MS, Finlay AY. The Dermatology Life Quality Index 1994–2007: a comprehensive review of validation data and clinical results. *Br J Dermatol* 2008; 159: 997–1035.
- Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI) – a simple practical measure for routine clinical use. *Clin Exp Dermatol* 1994; 19: 210–216.
- Rabin R, Charro F de. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001; 33: 337–343.
- Yang Y, Brazier J, Longworth L. EQ-5D in skin conditions: an assessment of validity and responsiveness. *Eur J Health Econ* 2015; 16: 927–939.
- Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011; 20: 1727–1736.
- Augustin M, Lange S, Wenninger K, Seidenglanz K, Amon U, Zschocke I. Validation of a comprehensive Freiburg Life Quality Assessment (FLQA) core questionnaire and development of a threshold system. *Eur J Dermatol* 2004; 14: 107–113.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. COSMIN checklist manual: COSMIN; 2012. [accessed 13 June 2018]. Available from: <http://www.cosmin.nl/images/upload/files/COSMIN%20checklist%20manual%20v9.pdf>.
- Fayers PM, Hand DJ. Factor analysis, causal indicators and quality of life. *Qual Life Res* 1997; 6: 139–150.
- Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013; 22: 1889–1905.
- Liang MH. Evaluating measurement responsiveness. *J Rheumatol* 1995; 22: 1191–1192.
- Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes* 2006: 70–75.