



European Association of Urology



Platinum Priority – Neuro-urology

Editorial by Antonella Giannantoni on pp. 280–281 of this issue

The Bladder Pain/Interstitial Cystitis Symptom Score: Development, Validation, and Identification of a Cut Score

Louise Humphrey^{a,*}, Rob Arbuckle^a, Rob Moldwin^b, Jorgen Nordling^{c,d}, Joop P. van de Merwe^e, Juliette Meunier^f, Tim Crook^g, Lucy Abraham^g

^a Mapi Values, Adelphi Mill, Bollington, Cheshire SK10 5JB, United Kingdom; ^b Hofstra University School of Medicine, Pelvic Pain Treatment Center, The Arthur Smith Institute for Urology, North Shore-LIJ Healthcare System, New Hyde Park, NY 11040, USA; ^c University of Copenhagen, Copenhagen, Denmark; ^d Department of Urology, Herlev Hospital, Herlev, Denmark; ^e Department of Immunology and Internal Medicine, Erasmus MC, Rotterdam, The Netherlands; ^f Mapi Values, Lyon, France; ^g Pfizer Ltd, Walton Oaks, Walton on the Hill, Tadworth, Surrey KT20 7NS, United Kingdom

Article info

Article history:

Accepted October 6, 2011

Published online ahead of print on October 17, 2011

Keywords:

Bladder pain syndrome
Painful bladder syndrome
Interstitial cystitis
Screener
Qualitative research
Cut score
Receiver operating characteristic (ROC) curve
Questionnaire

Abstract

Background: There is a need to develop a self-report measure that reliably identifies moderate to severe bladder pain syndrome (BPS) patients for inclusion into clinical trials to assess the efficacy of new BPS treatments.

Objective: To develop and validate a patient-reported symptom-based instrument, the Bladder Pain/Interstitial Cystitis Symptom Score (BPIC-SS), for clinical trial eligibility of BPS patients.

Design, setting, and participants: Stage 1: Qualitative concept elicitation (CE) interviews were conducted with BPS patients in France ($n = 12$), Germany ($n = 12$), and the United States (US) ($n = 20$), and overactive bladder (OAB) ($n = 10$) patients in the US for comparison. Stage 2: Cognitive debriefing (CD) interviews were performed with US BPS patients ($n = 20$). Stage 3: An observational study with 99 BPS, 99 OAB, and 100 healthy participants in the US was used to perform item reduction, identify cut scores, and validate the measure. A cut score was defined using logistic regression and receiver operating characteristic curves. Psychometric properties, including test-retest reliability, were assessed.

Measurements: In addition to the BPIC-SS, the Pelvic Pain and Urgency/Frequency Patient Symptom Scale, the Interstitial Cystitis Symptom Index, a Clinician Global Impression of Severity, and a Patient Global Impression of Change were included in the observational study.

Results and limitations: In CE, reported symptoms were bladder pain, persistent urge to urinate, and high urinary frequency. In CD, 13 items were deleted, and 15 were retained. Based on validation analyses, qualitative findings, and clinical relevance, the instrument was reduced to eight items that had strong sensitivity (0.72) and specificity (0.86) with a cut score ≥ 19 to determine clinical trial inclusion. Psychometric properties were strong.

Conclusions: The BPIC-SS is a reliable, valid, and appropriate questionnaire to select BPS/interstitial cystitis patients for clinical trials.

© 2011 European Association of Urology. Published by Elsevier B.V. All rights reserved.

* Corresponding author. Mapi Values, Adelphi Mill, Bollington, Cheshire SK10 5JB, United Kingdom. Tel. +1 625 577245; Fax: +1 625 5787328.

E-mail address: Louise.Humphrey@mapivalues.com (L. Humphrey).

1. Introduction

Bladder pain syndrome (BPS), also referred to as *interstitial cystitis* (IC) or *painful bladder syndrome* [1,2], is a chronic

bladder condition characterized by bladder pain, increased urinary frequency, and urge to urinate [1]. Prevalence estimates vary from 67 to 230 per 100 000 women [3]; 5–10% of diagnosed patients are men [4,5].

BPS represents a high unmet medical need because there is a lack of effective treatments for this condition. Many challenges confront clinical trials for novel therapies. A key one is the identification of an appropriate population. Research criteria developed at the National Institute of Diabetes, Digestive and Kidney Diseases (NIDDK) IC Workshop [6], which includes “objective” evidence of disease indicated by glomerulations or Hunner ulcer, are widely used to diagnose BPS. However, these criteria were devised to define a homogeneous population for research rather than diagnosis [2]. Approximately two-thirds of patients whom experienced clinicians regard as definitely or very likely to have BPS would not meet the NIDDK criteria [7,8]. Additionally, the presence of glomerulations or ulcer has no relationship to symptom severity. Consequently, the urologic community has moved towards symptom-based diagnostic criteria for BPS (alongside exclusion of confusable diseases) [1,2], although this also presents challenges due to overlapping symptoms with other conditions. For example, frequency and urge to urinate are also part of the symptom complex for overactive bladder (OAB), a condition often confused with BPS.

BPS clinical trials for investigational medicines need to recruit a confirmed BPS population with moderate to severe symptom burden, to ensure efficient statistical design and adequate benefit-risk ratio. A daily patient-completed symptom diary during a 1- to 2-wk screening period may identify this population, but those not meeting the symptom criteria may be excluded from the trial, causing additional burden to patients and investigators. Thus it is more efficient to identify these patients at initial entry to the study using a 7-d recall self-report symptom measure prior to further screening using a diary.

A review of existing patient-completed measures (eg, Pelvic Pain and Urgency/Frequency Patient Symptom Scale [PUF] and the Interstitial Cystitis Symptom Index [ICSII]) concluded that existing measures do not meet current standards for the development of patient reported measures and there is a need to develop a new patient-friendly measure with good sensitivity and specificity [9–12]. Thus the aim of the present study was to develop a new measure of BPS symptoms that could be used to screen patients into trials. The measure has been developed with patient and clinical input and using methods that meet standards for patient-completed measures [9–12].

2. Patients and methods

2.1. Inclusion and exclusion criteria

For all stages, eligible BPS patients had to have received a urologist-confirmed diagnosis of BPS with exclusion of confusable diseases (eg, OAB, endometriosis, or cervical, uterine, and ovarian cancer); experienced chronic pelvic pain (>6 mo); and reported pressure or discomfort related to the urinary bladder and one or more other urinary symptom. BPS patients had to have had a cystoscopy within 2 yr to confirm absence of other significant lower urinary tract pathology and to assess for cystoscopic features of BPS, although the presence of BPS features was not required for inclusion. Patients with significant physical or psychiatric comorbidities, confusable conditions, or passive urinary incontinence were excluded.

In stage 1, OAB patients provided information to inform how BPS symptoms differ from OAB symptoms, the most relevant confusable condition. In stage 3, OAB patients were included to evaluate the specificity and sensitivity of the measure for distinguishing between confusable conditions. In both stages, OAB patients had to have a physician-based diagnosis of OAB with symptoms ≥ 6 mo. OAB patients with any neurologic condition, pelvic organ prolapse, urinary tract infection (≤ 6 wk), or a history of BPS or bladder pain were excluded. Of note, it is not generally necessary for patients to be treatment free or under steady treatment for patient-reported outcome (PRO) development [9]. Healthy controls (HCs) in stage 3 were ≥ 18 yr of age, with no self-reported history of BPS or OAB (not physician verified). All participants provided written informed consent.

2.2. Stage 1: Concept elicitation and item generation

Forty-four exploratory interviews were conducted with BPS patients in France ($n = 12$), Germany ($n = 12$), and the US ($n = 20$) to gather information about experiences of BPS symptoms. Ten OAB patients in the US were interviewed to understand how BPS differs from OAB, particularly with regard to the urge to urinate. Concept elicitation (CE) interviews included open-ended questions with direct follow-up probes if key topics were not mentioned. Interviews in all stages were audiotaped and transcribed verbatim. Qualitative analysis was conducted in the original language by a native speaker. Using grounded theory methods [13] and ATLAS.ti software [14], quotes were assigned a code determined by the underlying concept and grouped into higher level concepts. Conceptual saturation was assessed [15].

Following analysis, researchers from all countries met to generate questionnaire items using a conceptual framework developed from the qualitative data. Clinical BPS experts were present to ensure the items were relevant and no clinically important symptoms were missed (authors RM, JN, and JM).

2.3. Stage 2: Cognitive debriefing

Twenty BPS patients in the US were administered the questionnaire and asked detailed questions about comprehension and relevance to ensure it adequately measured the concepts and items were understood and correctly interpreted. As per CE, analysis was conducted using ATLAS.ti software [14].

2.4. Stage 3: Development of cut scores and psychometric validation

An observational noninterventional study was conducted in the US with 300 participants (100 with BPS, 100 with OAB, and 100 HCs) to identify a cut score to differentiate between BPS and non-BPS patients and for psychometric validation. The newly developed Bladder Pain/Interstitial Cystitis Symptom Score (BPIC-SS), ICSI [16], and PUF [17], as well two single-item measures (Clinician Global Impression of Severity [CGI-S] and Patient Global Impression of Change [PGI-C]), were mailed to participants for completion. All participants completed the questionnaires at baseline, and the first 50% in each patient group who returned their questionnaires were mailed a second package for completion 7–14 d later.

The sample was divided into a test sample ($n = 150$) for item reduction and scoring, and a validation sample ($n = 150$) for validation of the scores. The test-retest sample included patients in the BPIC-SS baseline sample who completed one or more BPIC-SS item at follow-up and reported “no change” on the PGI-C ($n = 89$). Table 1 details the analyses performed to determine the structure of the BPIC-SS

Table 1 – Statistical methods used in stage 3 to determine structure and cut scores of the Bladder Pain/Interstitial Cystitis Symptom Score and to perform the psychometric validation of the final version

Method	Details of test
Item response distributions	Frequency and percentage of response choices including missing data were described at baseline for each of the BPIC-SS items per group and total sample. Items that showed unfavorable response distributions (heavily skewed or bimodal distributions, floor/ceiling effects) were taken into account when deleting items.
EFA	EFA (including interitem correlations) was conducted using varimax rotation and included test sample patients who completed all questionnaire items. This was used to aid in the development of item-scale structures and to explore item groupings. Due to the heterogeneity of the items, it was recognized a priori that a strong scale structure might not emerge. Findings were interpreted in light of the qualitative findings and clinical relevance of items.
ULR	To evaluate the extent each item was individually associated with each diagnosis, ULR was conducted with the diagnosis (BPS, OAB, or HC) as the explained variable and each BPIC-SS item as a covariate. Items were ordered according to their ability to discriminate between BPS patients and other participants. Items with low discriminant accuracy were considered for deletion. In addition, ULR was also conducted with diagnosis as the explained variable and each BPIC-SS score (created following item reduction) as a covariate.
ROC curves	ROC curves were plotted for each BPIC-SS score and the AUC was calculated to evaluate its discriminant properties. If a score has to discriminate between two groups, the corresponding ROC curve is a plot of sensitivity as a function of 1-specificity for the possible cut-off points. The wider the arc of the curve, the greater the AUC, and the more discriminative the score. An AUC of 1.00 represents a score that correctly discriminates on 100% of occasions. The Youden index indicates the highest value for the optimal cut-off score, which maximizes sensitivity and specificity.
Test-retest reliability	Reliability refers to the degree to which an instrument is free from measurement error and yields the same scores each time it is administered, all other things being equal [20–22]. Test-retest reliability of the BPIC-SS was evaluated using a signed rank and by calculating intraclass correlation coefficients, the concordance correlation coefficient, and Pearson correlations to evaluate changes in BPIC-SS scores between baseline and follow-up in patients who scored “no change” on the PGI-C. Coefficients >0.70 were considered evidence of acceptable test-retest reliability [20–22].
Concurrent validity	Correlations between the BPIC-SS and validated instruments measuring similar/related concepts were examined. It was expected that scales measuring more similar concepts would correlate more highly than domains that measure less similar concepts [20]. Correlations between the scores of the BPIC-SS and the scores of the PUF and ICSI were examined.
Clinical validity	Clinical validity refers to the ability of a patient completed instrument to detect variability among patients with different clinical severity levels [20]. The clinical validity of the BPIC-SS was explored by conducting a test where comparisons between the mean BPIC-SS scores for patients of different severity levels according to their scores on a CGI-S (completed by the patients' physician during screening) were undertaken. It was expected that patients rated as having more severe BPS (according to CGI-S) would score higher/worse on the BPIC-SS, with statistically significant differences among the groups.

BPIC-SS = Bladder Pain/Interstitial Cystitis Symptom Score; BPS = bladder pain syndrome; OAB = overactive bladder; HC = healthy control; PUF = Pelvic Pain and Urgency/Frequency Patient Symptom Scale; ICSI = Interstitial Cystitis Symptom Index; CGI-S = Clinician Global Impression of Severity; ULR = univariate logistic regression; EFA = exploratory factor analysis; ROC = receiver operating characteristic; AUC = area under the curve.

and identify a cut score. In addition, clinical relevance or importance and previous qualitative research were considered during item reduction.

2.5. Ethics

Ethical oversight for all stages was provided by Copernicus, a centralized independent review board in the US. For stage 1, ethics approval was obtained in Germany from the Freiburger Ethik Kommission but was not required for qualitative research in France.

3. Results

3.1. Demographic and clinical characteristics for concept elicitation and cognitive debriefing interviews

Table 2 details the demographic and clinical characteristics for the CE and cognitive debriefing (CD) samples. In all samples the majority were female and white; mean age was 48 yr in the CE sample and 43 yr in the CD sample. All characteristics were similar across the country samples except that the German BPS patients were older (mean: 57 yr) and had been diagnosed for a longer period of time (mean: 89 mo).

3.2. Stage 1: Concept elicitation results and item generation

During the CE interviews, BPS patients most frequently talked about experiencing three core symptoms: bladder pain, urge to urinate, and high urinary frequency. BPS patients also talked about bladder discomfort and pressure, difficulty initiating urination, a low or variable volume of urine, and urinary leakage. Examples of quotes are listed in Table 3. Bladder pain was the clearest differentiator between BPS and OAB; almost all BPS patients reported pain compared with a few OAB patients, and where pain was described by OAB patients, it was more related to pressure feelings. IC patients talked about “urge to urinate” in the context of avoiding or relieving bladder pain, whereas OAB patients tended to talk about urge in terms of a sudden and immediate need and avoiding urinary leakage. Saturation was achieved within all country samples and in the overall sample. No country-level differences in the descriptions of symptoms were found.

As a result, a conceptual framework [9] was developed and used to generate items to measure the core BPS symptom concepts of bladder pain/pressure/discomfort (5 items), urge to urinate (15 items), and urinary frequency (8 items). A 7-d recall period was chosen to ensure a long enough period to get a sense of the patient's current disease

Table 2 – Demographic and clinical characteristics of concept elicitation and cognitive debriefing samples (stages 1 and 2)

Demographic and clinical characteristic	CE stage					CD stage
	OAB (n = 10)	BPS US (n = 20)	BPS Ger (n = 12)	BPS Fr (n = 12)	BPS Total (n = 44)	BPS (n = 20)
Age of patient, yr						
Mean	47	45	57	47	48	43
Range	31–69	22–62	37–72	25–69	22–72	21–61
Gender, % (n)						
Male	20 (2)	25 (5)	25 (3)	17 (2)	23 (10)	20 (4)
Female	80 (8)	75 (15)	75 (9)	83 (10)	77 (34)	80 (16)
Ethnicity, % (n)						
Hispanic	0	5 (1)	0	NA*	3 [†] (1)	0
White	70 (7)	65 (13)	100 (12)		78 (25)	75 (15)
African American	20 (2)	25 (5)	0		16 (5)	20 (4)
Asian	10 (1)	0	0		0	0
Filipino	0	5 (1)	0		3 (1)	0
Missing						5 (1)
Highest education level, % (n)						
Some high school	10 (1)	5 (1)	0	0	2 (1)	0
High school diploma	30 (3)	20 (4)	0	0	10 (4)	30 (6)
Some years of college	40 (4)	20 (4)	8 (1)	0	11 (5)	20 (4)
Certificate program	0	0	84 (10)	0	23 (10)	5 (1)
Beauty diploma	0	0	0	8 (1)	2 (1)	0
Community diploma	0	0	0	8 (1)	2 (1)	0
College or university degree	20 (2)	45 (9)	8 (1)	67 (8)	41 (18)	30 (6)
Graduate or professional degree	0	10 (2)	0	17 (2)	9 (4)	15 (3)
Months since diagnosed						
Mean	20	28	89	24	44	57
Range	5–38	2–119	2–240	1–95	1–240	9–243
Missing data	(1)					
Years since first symptomatic						
Mean	2	8	12	6	8	8
Median	2	4	8	4	4	6
Min, max	0, 5	0, 40	2, 35	1, 24	0, 40	0, 21
Missing data	(1)	(2)			(2)	(1)
Patients currently taking treatment for bladder condition [§] , % (n)						
Licensed pharmacologic	40 (4)	80 (16)	75 (9)	58 (7)	73 (32)	100 (20)
Off-label pharmacologic	10 (1)	45 (9)	50 (6)	67 (8)	52 (23)	45 (9)
Over-the-counter pharmacologic	10 (1)	15 (3)	25 (3)	8 (1)	16 (7)	35 (7)
Nonpharmacologic	70 (7)	5 (1)	58 (7)	25 (3)	25 (11)	50 (10)
PUF total score						
Mean	13	24	22	21	23	21
Range	5–21	10–31	13–33	12–27	10–33	11–34
ICSI total score						
Mean	11	15	14	14	14	13
Range	6–16	4–19	9–20	8–20	4–20	7–18
Has a cystoscopy been performed? % (n)						
Yes	† NA	† NA	† NA	† NA	† NA	65 (13)
No						30 (6)
Missing data						5 (1)
Has a hydrodistention been performed? % (n)						
Yes	–	–	–	–	–	60 (12)
No						35 (7)
Missing data						5 (1)
Features, % (n)						
Tears/cracks						15 (3)
Scar						10 (2)
Glomerulations						35 (7)
Mucosal bleeding/blood effluent						30 (6)
Erythematous patches						5 (1)
Hunner lesions						5 (1)
None						0
Missing data						5 (1)

CE = concept elicitation; CD = cognitive debriefing; OAB = overactive bladder; BPS = bladder pain syndrome; US = United States; GER = Germany; FR = France; NA = not applicable; PUF = Pelvic Pain and Urgency/Frequency Patient Symptom Scale; ICSI = Interstitial Cystitis Symptom Index.

* Ethnicity data were not obtained in France as per local regulations.

[†] Total n is 32 because ethnicity data were not provided for French patients.

[‡] In CE (stage 1), cystoscopy was a requirement for study eligibility, but surgery results were not collected, so these data are not available.

[§] Patients could select all treatments that applied; therefore total may sum to >100%.

Table 3 – Examples of patient quotes for symptom concepts identified in concept elicitation interviews (stage 1)

Symptom concept	Subconcept	Example of BPS patient comments
Pain/pressure/discomfort perceived to be associated with the bladder	Bladder pain	<i>It feels like there is somebody in my body kicking the hell out of my bladder. That's what it feels like. Woman, age 31</i>
	Bladder discomfort	<i>It's a feeling that's always there. Oh God, how do I tell you? Like I just went to the bathroom and I still feel the discomfort. So although I cleared my bladder, I still feel discomfort there. Woman, age 34</i>
	Bladder pressure	<i>Yes, I can go after a nap, but I don't have any pain. Just some pressure, but that's all. Man, age 50</i>
Urinary frequency	Daytime urinary frequency	<i>Depending on good day, bad day. There's a lot of variables there. But I guess on a bad day, maybe 60 times a day. Woman, age 35</i>
	Nighttime urinary frequency	<i>Unfortunately, it's that it keeps happening. Last night, for instance, I had to get up four or five times. The night before, I don't know, maybe more. It's very irregular. When I get up two or three times, I'm happy. Man, age 55</i>
	Nocturia	<i>But last night I maybe woke up once or twice to go the bathroom, which is very good for me because there can be times where every 10 minutes I'm going. I can't even close my eyes. I just go to the bathroom and come back. I know I'm going to come right back. It's always worse at night. Woman, age 37</i>
	Bother associated with daytime urinary frequency	<i>I think that I can probably go to the bathroom 30 times a day, sometimes more, sometimes less. But I almost think sometimes now it's becoming psychological. I think it turns—can almost turn—into a psychological thing. It makes you feel like you're nuts is what I'm saying. When I say psychological, it makes you feel like you're losing your mind. Woman, age 44</i>
	Bother associated with nighttime urinary frequency Bother associated with nocturia	<i>And then nighttime's the worst for me. I would say I can deal with the daytime stuff. It's annoying and frustrating and very inconvenient. Woman, age 27 That's another reason why I'm up all night, because the pain'll wake me up because I've drank so much. In the evening, I drink all evening long. I'll probably drink 32 ounces of water from maybe 5:00 until 9:00. And then I pay for it because the pain wakes me up to have to go the bathroom because my bladder is really bulging then. Woman, age 59</i>
Urge to urinate	Persistent urge to urinate	<i>Probably the most frustrating part is I'll go to the restroom, relieve myself, and then sometimes within 5 minutes I have to go right back and try again. Woman, age 27</i>
	Sudden and immediate urge to urinate The necessity to urinate driven by bladder pain/discomfort/pressure	<i>Because you got to go now, you can't hold it in. I could hold it in forever before. Even when I had to go, I had control over that muscle group. I don't anymore. Man, age 22 Well, the pain is the urgency. Once I got that pain, it means that I've got to go. So the pain is a warning sign for me, and it's not like you just get a full bladder and then you go. Well, it's the pain that's telling me that I need to go. Man, age 51</i>
Difficulty initiating urination	NA	<i>I always have to push. I always have to push a little bit until the urine really comes out. It doesn't just start running all on its own. I always have to push a little bit, always push a little bit. Woman, age 68</i>
A low or variable volume of urine	NA	<i>You feel like you could sit there and pee and pee and pee. You think in your head that's how much is in you. Whether or not that's true or not, only a little comes out. Sometimes nothing comes out. Woman, age 37</i>
Urinary leakage	NA	<i>I don't know if it's that I'm waiting for the last minute, trying to hold it off, I don't think that's it. It's just that the pain won't let me. I just got to let it go. Man, age 56</i>

BPS = bladder pain syndrome; NA = not applicable.

severity but not so long that there would be concerns about the accuracy of recall. Standard practice for instrument development is to take a large pool of items derived from CE and reduce this number based on the results of content and psychometric validation. Thus the intention was to test the 28-item BPIC-SS in CD and delete items based on patient feedback and psychometric analysis.

3.3. Stage 2: Cognitive debriefing

Patients found the 28-item BPIC-SS easy to understand and answer. Patients did not identify any missing symptoms of importance. The items and response scales were generally well understood and interpreted consistently and correctly. All patients understood and could use the 7-d recall. Some changes were recommended based on CD findings; almost all BPS patients felt the bladder pain and pressure questions measured two distinct concepts, but many reported that “bladder discomfort” overlapped with “pain” and “pressure.” It was concluded that “discomfort” is captured by the pain and pressure items, and so the discomfort questions were

deleted. Nocturia items were deleted because patients did not differentiate between those and the nighttime frequency items. Overall, 13 items were deleted, 10 were retained without changes, and 5 were retained with minor wording revisions. A 15-item BPIC-SS was taken into stage 3.

3.4. Stage 3: Development of cut scores and psychometric validation

Table 4 lists the characteristics of the 298 eligible patients included in stage 3. Two patients were enrolled but not included due to incomplete information on the clinician-completed form. Most were female (79.5%; n = 237) with a mean age of 47.5 yr, and most were white (74.5%; n = 222). Cystoscopic features of BPS were present in 76.8% (n = 76) of BPS patients.

3.4.1. Determination of the new Bladder Pain/Interstitial Cystitis Symptom Score structure

Missing data were very low for all items; no items were missed by more than three participants (1.01%; n = 3). For

Table 4 – Demographic and clinical characteristics of psychometric validation sample (stage 3)

Demographic and clinical characteristics	BPS (n = 99)	OAB (n = 99)	HC (n = 100)
Age, yr			
Mean (SD)	48 (13)	51 (15)	44 (15)
Median	48	49	47
Range	21–78	22–83	19–72
Gender, % (n)			
Female	88 (87)	81 (80)	70 (70)
Male	12 (12)	19 (19)	30 (30)
Ethnicity, % (n)			
Hispanic	3 (3)	11 (11)	8 (8)
White	80 (79)	73 (72)	71 (71)
African American	16 (16)	15 (15)	17 (17)
Other	1 (1)	1 (1)	4 (4)
Months since diagnosis			
Mean	48	24	NA
Median	24	12	
Min, max	0, 372	0, 168	
Years since first symptomatic			
Mean	8	3	NA
Median	5	2	
Range	1–41	0–16	
Received treatment for bladder condition, % (n) [§]			
Pharmacologic treatments	81 (80)	75 (74)	NA
Off-label treatments	58 (57)	6 (6)	
Over-the-counter pharmacologic	36 (36)	1 (1)	
Nonpharmacologic treatments	70 (69)	31 (31)	
PUF total score			
Mean	21	14	4
Range	7–35	2–31	0–22
ICSI total score			
Mean	13	9	3
Range	3–20	1–18	0–13
Type of cystoscopy, % (n)			
General anaesthetic	53 (52)	NA	NA
Awake	47 (47)		
Cystoscopic features of BPS were present, % (n)	77 (76)		
Bladder capacity at cystoscopy			
n	81		
Mean	393		
Median	320		
Range	60–1100		
Features present at cystoscopy, % (n)			
Tears/cracks	21 (21)		
Scar	18 (18)		
Glomerulations	35 (35)		
Mucosal bleeding, blood effluent	20 (20)		
Erythematous patches	26 (26)		
Hunner lesions	6 (6)		
None	14 (14)		
Not available	9 (9)		

BPS = bladder pain syndrome; OAB = overactive bladder; HC = healthy control; SD = standard deviation; NA = not applicable; ICSI = Interstitial Cystitis Symptom Index; PUF = Pelvic Pain and Urgency/Frequency Patient Symptom Scale.

[§] Patients could select all treatments that applied; therefore total may sum to >100%.

all items, mean scores were worse in the BPS sample, followed by OAB, and HCs scored the best. Item reduction was conducted with the aim of retaining the most discriminative items while also retaining at least one item for each concept included in the conceptual framework. Clinical relevance or importance was also considered. Seven items were deleted based on exploratory factor analysis, logistic regression analyses (odds ratios), and receiver operating characteristic (ROC) curves. The eight-item BPIC-SS that resulted from this process included one “persistent urge to urinate” item, two “necessity to urinate driven by

bladder pain” items, two “bother associated with daytime and nighttime frequency” items, and three “bladder pressure/pain” items (Table 5). The item scores are summed to create a total ranging from 0 to 38.

Items measuring “bother” associated with urinary frequency were retained in preference to items asking about urinary frequency because “bother” was considered more reflective of actual frequency (and severity) than asking patients to recall frequency over a week. Statistical analysis also indicated the bother items were more discriminative. A “sudden and immediate urge to urinate”

Table 5 – Conceptual framework for final eight-item Bladder Pain/Interstitial Cystitis Symptom Score^{*} detailing the items retained, the item groupings, and the scoring

Concept	Subconcept	Item
Urge to urinate	Persistent/constant need to urinate	In the past 7 days, how often did you still feel the need to urinate just after you urinated?
Urinary frequency	The necessity to urinate driven by bladder pain	In the past 7 days when you urinated, how often was it because of pain in your bladder?
	Bother associated with daytime urinary frequency	In the past 7 days, how often did you urinate to avoid pain in your bladder from getting worse?
Bladder pain and pressure	Bother associated with nighttime urinary frequency	In the past 7 days, how bothered were you by frequent urination during the daytime?
	NA	In the past 7 days, how bothered were you by having to get up during the night to urinate?
		In the past 7 days, how often did you have a feeling of pressure in your bladder?
		In the past 7 days, how often did you have pain in your bladder?
		Select the number that describes your worst bladder pain in the past 7 days.

NA = not applicable.

^{*} For the scoring, a single score is created by summing all eight items to create a total score ranging from 0 to 38.

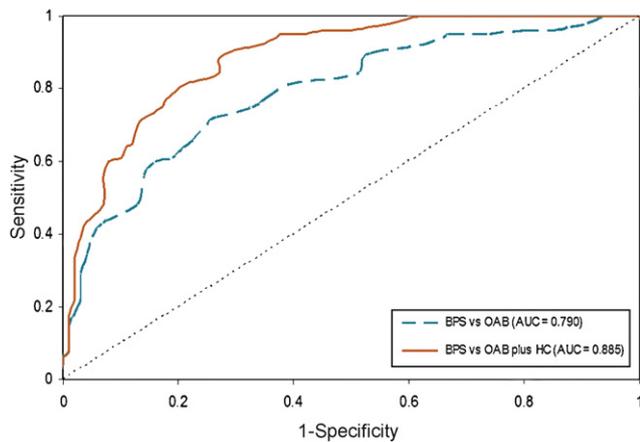


Fig. 1 – Receiver operating characteristic curves for the final eight-item Bladder Pain/Interstitial Cystitis Symptom Score in all samples. AUC = area under the curve; BPS = bladder pain syndrome; HC = health control; OAB = overactive bladder.

was reported by BPS patients in stage 1, but items measuring this did not discriminate between OAB and BPS so they were deleted; this symptom might be important to measure as an outcome in BPS but has little utility as a discriminator.

3.4.2. Determining the cut score of the new Bladder Pain/Interstitial Cystitis Symptom Score

To select a cut score for identification of BPS patients that maximized the sensitivity and specificity of the BPIC-SS, ROC

curves for the total score were generated in the different analysis samples (Fig. 1). Cut scores associated with the optimal Youden index ranged from 18 to 25. However, the cut score of 18 emerged most frequently from the different samples. A retrospective examination demonstrated that the cut score of 18 was acceptable across samples or analysis groups (Table 6). Therefore patients have to score ≥ 19 of 38 on the BPIC-SS to be eligible to enter a BPS trial.

3.4.3. Psychometric validation of the new Bladder Pain/Interstitial Cystitis Symptom Score

Table 7 shows the results of psychometric validation. The BPIC-SS demonstrated good test-retest reliability and strong concurrent and clinical validity. All a priori tests, thresholds, and expectations were met or surpassed in relevant patient groups.

4. Discussion

The BPIC-SS was developed using recognized methodology [9–12]. Expert BPS clinicians provided guidance to ensure clinical relevance of items. Items were generated from qualitative research with BPS patients in three countries to ensure the most important symptoms were included using patient-friendly, easily translated, non-culturally specific language. OAB patients were interviewed to ensure the measure would differentiate BPS patients from patients with similar symptoms. CD provided evidence that items were relevant and comprehensible. Logistic regressions suggested that many item combinations would provide an equally discriminative measure. Thus, as well as statistical results,

Table 6 – Results of receiver operating characteristic analysis for a cut score of 18 in all samples

Sample	Analysis	AUC	Youden index	Sensitivity	Specificity
Test sample	BPS vs OAB	0.766	0.44	0.67	0.77
	BPS vs OAB plus HC	0.878	0.57	0.72	0.85
Validation sample	BPS vs OAB	0.810	0.46	0.87	0.60
	BPS vs OAB plus HC	0.891	0.63	0.87	0.77
Total sample	BPS vs OAB	0.790	0.45	0.72	0.74
	BPS vs OAB plus HC	0.885	0.58	0.72	0.86

AUC = area under the curve; BPS = bladder pain syndrome; OAB = overactive bladder; HC = healthy control.

Table 7 – Psychometric validation results of the BPIC-SS

Validation criteria	Sample	Test statistics			
		BPIC-SS and ICSI		BPIC-SS and PUF	
		Pearson correlation coefficient	Kappa coefficient [95% CI]	Pearson correlation coefficient	Kappa coefficient [95% CI]
Concurrent validity	BPS (N = 99)	0.66	0.30 [0.09–0.50]	0.80	0.45 [0.24–0.66]
	OAB (N = 95)	0.62	0.26 [0.14–0.38]	0.85	0.59 [0.44–0.74]
	HC (N = 99)	0.71	0.31 [–0.02–0.64]	0.78	0.80 [0.40–1.00]
	Total (N = 293)	0.84	0.55 [0.47–0.63]	0.92	0.74 [0.66–0.81]
Test-retest reliability		Mean BPIC-SS total score (SD)		p-value (signed rank)	Intraclass correlation coefficient (ICC)
	Test-retest (N = 89)	Baseline	Follow-up	0.9465	0.95
	11.2 (11.6)	11.0 (12.0)			
Clinical validity		Mean BPIC-SS total score (SD)			p-value
	Total (N = 194)	Very severe/Severe (N = 49)	Moderate (N = 89)	Mild/Very mild (N = 56)	<0.0001
	28.0 (7.8)	17.7 (9.3)	15.6 (9.2)		

item selection took into account the qualitative findings, clinical importance, and need to identify a moderate to severe population for clinical trials. Hence the final grouping included all key symptom concepts but more bladder pain items (the defining symptom of BPS [6]), and the items that were most discriminative between BPS and OAB. The resulting BPIC-SS has been developed to the standards set within the guidance of the US Food and Drug Administration (FDA) for PRO development and has strong sensitivity and specificity, making it appropriate for assessing the eligibility of BPS patients for clinical trials. Although developed from patient interviews in three countries, further validation should be performed in other languages to confirm cross-cultural validity. The BPIC-SS has been translated into 34 languages and included in a multicountry clinical trial that could be used for further validation.

In stage 3 of the present study we provide evidence that the BPIC-SS can discriminate BPS patients from OAB patients and HCs. Further testing of the ability of the tool to discriminate BPS patients from those with other confusable conditions (eg, endometriosis, urinary tract infections) would be a valuable avenue for future research.

Another way to consider the discriminative power of the BPIC-SS is compare it with that of the PUF when administered in this study. For the PUF, a cut score ≥ 13 was suggested as defining a moderate to severe symptom burden [18]. For the PUF, 53% of OAB patients had a score ≥ 13 and therefore would be (incorrectly) categorized as BPS patients. In contrast, only 40% of OAB patients would be classed as BPS based on their BPIC-SS scores. The specificity of the BPIC-SS increases to 0.86 when HCs are included in the analysis, again providing support that the more distinct the comparison groups, the more discriminative the tool. For the ICSI, a score ≥ 7 (as recommended by the developers [16]) correctly identified 61–77% of patients with BPS in a managed care population [5]. In the current study, 93% of BPS patients scored ≥ 7 ; however, 91% of OAB patients also scored ≥ 7 . In summary, the BPIC-SS appears more discriminative than the PUF or the ICSI and has been

developed to the standards set within the FDA's guidance for PRO development [9].

Potential limitations of the study were the definitions of BPS and OAB formulated by the entry criteria and physician diagnosis. The BPS entry criteria reflected clinical practice diagnostic guidelines rather than the stricter research criteria of the NIDDK [6]. However, the aim was to include patients similar to more recently defined clinical trial populations, to whom the measure will be administered. The entry criteria were therefore broader than NIDDK while still ensuring that patients with comorbid conditions were excluded. Similarly, the OAB diagnosis was confirmed by physicians and not using objective criteria such as urodynamics. However, the OAB entry criteria were consistent with ICS guidelines [19]. It is possible that a urodynamically defined population may not have been representative of the symptomatically defined OAB population, from which the BPIC-SS should help discriminate BPS patients.

Although developed for the purpose of screening into trials, the BPIC-SS could also be used for measuring outcome or treatment effects during a clinical study. However, the qualitative data reported here were also used to develop the BPIC-eDiary, a daily electronic measure of BPS symptoms. We recommend that the BPIC-eDiary is likely to have stronger validity, reliability, and ability to detect changes in symptoms over time, and so for outcome measurement it should be used in preference to the BPIC-SS.

5. Conclusions

The BPIC-SS is a reliable, valid, and discriminative clinical trial eligibility tool that assists in the identification of moderate to severe BPS patients.

Note: The BPIC-SS is available free of charge via the PROLUTS Web site: <http://www.prolutssh.com/index.html>.

Author contributions: Louise Humphrey had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Humphrey, Arbuckle, Moldwin, Nordling, van de Merwe, Meunier, Crook, Abraham.

Acquisition of data: Humphrey, Arbuckle.

Analysis and interpretation of data: Humphrey, Arbuckle, Moldwin, Nordling, van de Merwe, Meunier, Crook, Abraham.

Drafting of the manuscript: Humphrey, Arbuckle, Crook, Abraham.

Critical revision of the manuscript for important intellectual content: Humphrey, Arbuckle, Moldwin, Nordling, van de Merwe, Meunier, Crook, Abraham.

Statistical analysis: Meunier.

Obtaining funding: Humphrey, Arbuckle, Crook, Abraham.

Administrative, technical, or material support: Humphrey.

Supervision: Crook, Abraham.

Other (specify): None.

Financial disclosures: I certify that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/ affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Louise Humphrey, Rob Arbuckle, and Juliette Meunier were employees of Mapi Values, a consultant paid by Pfizer to perform the study and develop the manuscript.

Funding/Support and role of the sponsor: Pfizer helped interpret the data and review and approve the manuscript.

Acknowledgment statement: The authors would like to thank acknowledge the individuals with interstitial cystitis and overactive bladder syndrome and the healthy participants who took part in the study as well as Barbara Gordon from the Interstitial Cystitis Association and Jane Meijlink from the International Painful Bladder Foundation for their thoughts and valuable input throughout the study. The authors would also like to acknowledge the contributions of Sabine Bielfeldt, Kristina Fitzgerald, and Carla Dias-Barbosa in conducting the interviews. They would also like to thank Nicola Barnes and Nicola Bonner for their input into the qualitative analysis. Finally, we would like to acknowledge the involvement of Tara Symonds, David Scholfield, Ian Mills, and Benoit Arnould in the item-generation meeting and for guidance and scientific advice throughout the project. We also acknowledge guidance from Muriel Viala-Danten related to the item-reduction analyses.

References

- [1] Macdiarmid SA, Sand PK. Diagnosis of interstitial cystitis/painful bladder syndrome in patients with overactive bladder symptoms. *Rev Urol* 2007;9:9–16.
- [2] van de Merwe JP, Nordling J, Bouchelouche P. Diagnostic criteria, classification, and nomenclature for painful bladder syndrome/interstitial cystitis. An ESSIC proposal. *Eur Urol* 2008;53:60–7.
- [3] Forrest JB, Mishell DR. Diagnosis and management for interstitial cystitis/painful bladder syndrome. *Clin Courier* 2006;24:1–8.
- [4] Teichman JM, Parsons CL. Contemporary clinical presentation of interstitial cystitis. *Urology* 2007;69(Suppl):41–7.
- [5] Clemens JQ, Meenan RT, O’Keeffe Rosetti MC, Brown SO, Gao SY, Calhoun EA. Prevalence of interstitial cystitis symptoms in a managed care population. *J Urol* 2005;174:576–80.
- [6] National Institute of Diabetes, Digestive and Kidney Diseases. Interstitial cystitis. NIH publication 94–3220. Rockville, MD: US Public Health Services, National Institute of Health; 1994.
- [7] Nickel JC. Diagnosis of interstitial cystitis: another look. *Rev Urol* 2000;2:167.
- [8] Brewer ME, White WM, Klein FA, Klein LM, Waters WB. Validity of pelvic pain, urgency, and frequency questionnaire in patients with interstitial cystitis/painful bladder syndrome. *Urology* 2007;70:646–9.
- [9] US Department of Health and Human Services; FDA Center for Drug Evaluation and Research. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. International Society for Pharmacoeconomics and Outcomes Research Web site. <http://ispor.org/workpaper/FDA%20PRO%20Guidance.pdf>.
- [10] European Medicines Agency. Reflection paper on the regulatory guidance for the use of health related quality of life (HRQL) measures in the evaluation of medicinal products. International Society for Pharmacoeconomics and Outcomes Research Web site. <http://ispor.org/workpaper/EMA-HRQL-Guidance.pdf>.
- [11] Lasch KE, Marquis P, Vigneux M, et al. PRO development: rigorous qualitative research as the crucial foundation. *Qual Life Res* 2010;19:1087–96.
- [12] Leidy NK, Vernon M. Perspectives on patient-reported outcomes: content validity and qualitative research in a changing clinical trial environment. *Pharmacoeconomics* 2008;26:363–70.
- [13] Glaser BG, Strauss AL. The discovery of grounded theory. New York, NY: Aldine; 1967.
- [14] ATLAS.ti v.5.2 [computer program]. Berlin, Germany: Scientific Software Development GmbH.
- [15] Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 2006;18:59–82.
- [16] O’Leary MP, Sant GR, Fowler Jr FJ, Whitmore KE, Spolarich-Kroll J. The interstitial cystitis symptom index and problem index. *Urology* 1997;49(Suppl):58–63.
- [17] Parsons CL, Dell J, Stanford EJ, et al. Increased prevalence of interstitial cystitis: previously unrecognized urologic and gynecologic cases identified using a new symptom questionnaire and intravesical potassium sensitivity. *Urology* 2002;60:573–8.
- [18] Kushner L, Moldwin RM. Efficiency of questionnaires used to screen for interstitial cystitis. *J Urol* 2006;176:587–92.
- [19] Abrams P, Cardozo L, Fall M, et al. The standardisation of terminology in lower urinary tract function: report from the standardisation sub-committee of the International Continence Society. *Urology* 2003;61:37–49.
- [20] Hays RD, Anderson R, Revicki D. Assessing reliability and validity of measurement in clinical trials. *Quality of Life assessment in clinical trials: methods and practice*. Oxford University Press 1998;169–82.
- [21] Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality of life instruments: attributes and review criteria. *Qual Life Res* 2002;2:441–9.
- [22] Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159–74.