

## ORIGINAL ARTICLE

# Development of the Patient Harvey Bradshaw Index and a Comparison With a Clinician-based Harvey Bradshaw Index Assessment of Crohn's Disease Activity

Floor Bennebroek Evertsz', Msc,\* Céline C. M. Q. Hoeks, Msc,\* Pythia T. Nieuwkerk, PhD,\*  
Pieter C. F. Stokkers, PhD,† Cyriel Y. Ponsioen, MD, PhD,‡ Claudi L. H. Bockting, PhD,§  
Robbert Sanderma, PhD,|| and Mirjam A. G. Sprangers, PhD\*

**Goals and Background:** The objective is to develop a patient-based Harvey Bradshaw Index (P-HBI) of Crohn's Disease (CD) activity and to compare it with the clinician-based HBI of CD activity in CD outpatients.

**Study:** Consecutive patients with CD randomly completed the P-HBI either before or after the consultation. The gastroenterologist assessed patient's CD activity on the same day. Overall agreement between HBI and P-HBI was calculated with Spearman's  $\rho$  and Mann-Whitney  $U$  test. Agreement regarding active disease versus remission and agreement at item level was calculated by percent agreement and Cohen's  $\kappa$ .

**Results:** One hundred eighty-one (response rate 88.3%) CD patients participated. P-HBI and HBI showed a large correlation ( $r_s = 0.82$ ). The medians (interquartile range) of the total HBI (2; 0 to 4) and P-HBI (4; 1 to 7) were statistically significantly different ( $z = -8.411$ ;  $P < 0.001$ ). Fortunately, in 82.6% of the cases this difference between clinicians and patients was not clinically significant ( $< 3.2$ ). The percentage agreement between clinician and patient, judging CD as active or as in remission, was 77%,  $r_s = 0.56$ ,  $\kappa = 0.52$ , indicating a moderate agreement. P-HBI and HBI on frequent extraintestinal manifestations in CD varied from less than chance ( $\kappa = -0.02$ ) to a perfect agreement ( $\kappa = 1$ ). Patients tended to report more symptoms while completing the patient-

based questionnaire compared to what they reported to the clinician during consultation.

**Conclusions:** The P-HBI is the first step in developing a potential promising tool given its adequate agreement with the original HBI and its feasibility, especially in patients with low scores. Future research is necessary to develop a validated patient-based version studied in several patient populations.

**Key Words:** Crohn's disease, disease activity, Harvey Bradshaw index, inflammatory bowel disease validation.

(*J Clin Gastroenterol* 2013;47:850–856)

Crohn's disease (CD) is most often diagnosed at young age, has a chronic nature, and often requires life-long medical treatment. According to the 2010 International inflammatory bowel disease (IBD) task force meeting, incidence and prevalence of CD is steadily increasing worldwide.<sup>1–3</sup> The general course of CD is extremely unpredictable, characterized by periods of inflammation and remission. Disease flares can be intense and are frequently accompanied by increased pain, fatigue, and diarrhea. Clearly, CD can severely challenge the well-being of patients and limit their daily functioning.<sup>4</sup> As relapse is unpredictable and frequent, with a quarter to half of patients relapsing annually,<sup>5,6</sup> monitoring of disease activity is of vital importance.

In daily clinical practice, the treating clinician assesses disease activity in patients with CD. Frequently, clinicians use clinical activity indices, such as CD Activity Index (CDAI), Cape Town Index, Van Hees Index, or Harvey Bradshaw Index (HBI), as these are essential instruments in monitoring the course of CD.<sup>7</sup> Currently, the CDAI is the most widely validated index and is perceived as the golden standard. However, the index is complex, rather impractical, and time consuming.<sup>8</sup> Therefore, a simplified version of the CDAI, the HBI, is also commonly used (Appendix I, <http://links.lww.com/JCG/A66>).<sup>9–11</sup> Although some prefer the CDAI over its simplified version,<sup>9,10</sup> studies indicate that results from the HBI correlate well with those from the CDAI, with correlations ranging from 0.80 to 0.93.<sup>11,12</sup>

Knowing patients' CD activity is of great importance in clinical practice as well as medical research. Two studies have thus far developed a shortened assessment of CD activity derived from the CDAI<sup>13,14</sup> with the aim of easing the use of the CDAI in research. However, these short indices still require that patients record and sum up their disease activity for 7 consecutive days using a disease activity diary. Moreover, these indices are not translated into direct patient-based questionnaires, the requested

Received for publication September 17, 2012; accepted February 1, 2013.

From the Departments of \*Medical Psychology; ‡Gastroenterology, Academic Medical Center, Amsterdam; †Department of Gastroenterology, Saint Lucas Andreas Hospital, Amsterdam; §Department of Clinical Psychology, University of Groningen; and ||Health Psychology Section, Department of Health Sciences, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

F.B.E.: Conception and design of the study, development and translation of the P-HBI, statistical analysis, data interpretation, writing of first draft, and critical revision of the manuscript. C.C.M.Q.H.: Support in the literature searches, reference preparation, statistical analysis, data interpretation, and critical revision of the manuscript. P.T.N.: Statistical analysis, data interpretation, and critical revision of the manuscript. P.C.F.S.: Conception and design of the study, development and translation of the P-HBI. C.Y.P.: Data collection, interpretation, and critical revision of the manuscript. C.L.H.B.: Data interpretation and critical revision of the manuscript. R.S.: Critical revision of the manuscript. M.A.G.S.: Conception and design of the study, development and translation of the P-HBI, data interpretation, and critical revision of the manuscript.

The authors declare that they have nothing to disclose.

Reprints: Floor Bennebroek Evertsz', Msc, Department of Medical Psychology, Academic Medical Center, Meibergdreef 15, Amsterdam 1105 AZ, the Netherlands (e-mail: [f.bennebroek@amc.uva.nl](mailto:f.bennebroek@amc.uva.nl)).

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Website, [www.jcge.com](http://www.jcge.com).

Copyright © 2013 by Lippincott Williams & Wilkins

information is not well explained and the indices lack clear general instructions. These restraints strongly limit their use for medical research.

Hence, the first aim of this study was to develop an easy and undemanding patient-based questionnaire, which measures CD activity and can be completed by patients themselves. Therefore, we decided to use the HBI, as this simpler index can be easily translated into a patient-based questionnaire (P-HBI). As far as we know, this is the first study that develops a patient-based questionnaire of the HBI. As the utility of patient-assessed CD activity is dependent upon its agreement with the clinician-based assessment, our second aim was to assess agreement between the patient-based questionnaire, derived from the HBI, with the original clinician-based HBI. Ultimately, patient-based assessment of CD activity could increase the chance of early detection of relapse and could enable the investigation of CD on a larger scale.

## METHODS

### Study Population and Procedure

Consecutive patients with confirmed CD attending the IBD outpatient clinic of the Academic Medical Centre Amsterdam, from April to December 2010, were invited to participate in the study. Patients with insufficient command of Dutch were excluded. Participants were asked to complete the patient-modified HBI (P-HBI) in the hospital. To avoid order effects a random half of the patients completed the questionnaire before the outpatient consultation, while the other half completed the questionnaire after the consultation, wherever the consultation schedules permitted. Three clinicians participated in the study, blinded for the patients' responses. They assessed CD activity during the outpatient consultation by scoring the original HBI.

### Measures

#### Clinician-based HBI

The clinicians completed the Dutch version of the original HBI<sup>11</sup> (Appendix I, <http://links.lww.com/JCG/A66>). This questionnaire refers to disease symptoms during the previous week and includes 12 items. It is composed of 5 domains: general well-being; abdominal pain; number of liquid stools per day; abdominal mass (limited physical examination by the clinician) and extraintestinal manifestations of CD (eg, arthralgia, uveitis, erythema nodosum, aphthous ulcer, pyoderma gangrenosum, anal fissure, new fistula, and abscess). The questions concerning the 8 extraintestinal manifestations have a 2-fold "yes" or "no" option. We added 1 instruction to the item "number of liquid stools," that is, "In case of a stoma, please score deviation in comparison to its normal state." After recoding, the clinician-based HBI is able to categorize 2 types of patients: patients with inactive disease (HBI score < 5) and patients with active disease (HBI score  $\geq$  5).<sup>10,12</sup>

#### P-HBI

For patients, the Dutch version of the original HBI was translated into a patient-based questionnaire (Appendix II, <http://links.lww.com/JCG/A67>). This P-HBI was devised by 2 medical psychologists, 1 research assistant, and 1 gastroenterologist. All items within the P-HBI refer to symptoms during the previous week and medical terminology and disease symptoms were clarified. For example

"aphthous ulcers" is explained as "blister like spots in the mouth" and "a fissure" is described as "anal pain during bowel movements. Here we are referring to pain in the anus during the passage of stools." The first item about general well-being is formulated as follows: "If you would have to rate your general well-being by giving it a number, what number would you choose?" (1 = terrible, 10 = very well). The second item about abdominal pain is: "To what extent did you have abdominal pain during the previous week? Please indicate this on a scale from 0 to 10, where 0 stands for no abdominal pain and 10 for the worst imaginable abdominal pain." The third item on physical examination of abdominal mass was omitted, as we assumed that patients cannot adequately physically examine themselves. As a consequence, the patient-modified HBI consists of 11 items, instead of the 12 items in the clinician-based HBI. The fourth item is the question concerning number of daily liquid stools: "How many times on average per day did you have diarrhea during the previous week? Diarrhea is liquid stool that cannot be picked up (for instance with a fork)." For the questions about the 8 well-known extraintestinal manifestations, the response options for the patient were 3-fold: "yes," "no," and "I do not know." This third response option was added, as patients may not know whether they have a specific manifestation or may be unfamiliar with its specific medical terminology.

#### Debriefing Questions

We examined the comprehensibility of the P-HBI during a pilot study (n = 3). In general, patients reported that the questions were clear and easy to complete. Two major improvements on the basis of patients' feedback were made. First, in the second question about "abdominal pain" the scaling ranged from 1 to 10, which we changed into a range from 0 to 10. Second, in the question about "the number of liquid stools daily" 2 optional boxes were added, which patients can tick when having either a stoma or a pouch, since these can greatly affect the stool. Patients are asked to indicate how often they changed a bag in case of a stoma or had excrement in case of a pouch on average per day during the previous week compared to a regular week.

#### Demographic and Clinical Characteristics

CD diagnosis, sex, birth date, year of diagnosis, presence of stoma (yes/no), number of operations associated with CD and presence (no vs.  $\geq$  1) of comorbidity unrelated to CD (ie, 12 other illnesses) were measured through self-report.

#### Statistical Analysis

Assuming that 30% of patients have an active disease and that agreement is 0.23, which is higher than chance ( $\kappa$  0.63 vs. 0.40), 150 patients were needed, with 80% power and a 2-sided  $\alpha$  of 0.05. We used standard descriptive statistics to summarize the sociodemographic and clinical characteristics of included patients. We examined agreement between HBI scores of the clinician and patient on the total sum score and per item.

As previously indicated, the P-HBI had a third response option from question 6 to 13, "I do not know." We analyzed the patients' response "I do not know" in 3 different manners, either by scoring it as "no," as "yes," or as missing. As differences between these 3 modes of analysis

were negligible (data not shown), we decided to consider “I do not know” as “no” for further analysis.

### Patient-Physician Agreement at Total Sum Level

First, we assessed the strength of the correlation between the total HBI score assessed by the clinician and by the patient using Spearman's  $\rho$ . Correlation coefficients were interpreted as small ( $< 0.3$ ), medium (0.3 to 0.5), or large ( $> 0.5$ ). Second, we compared the total HBI score assessed by the clinician and by the patient using the Mann-Whitney  $U$  test. Analyses were performed with and without the item “abdominal mass.” Third, a difference of 3.2 points on the total HBI score was in a previous study considered to be clinically significant.<sup>12</sup> This means that the difference of 3.2 points is the smallest difference between the total scores of patients and clinicians that is considered to be clinically important. Therefore, we considered a difference  $\geq 3.2$  points between the HBI score assessed by the clinician and by the patient as a clinically significant difference, and a difference  $< 3.2$  points as not clinically significant. We thus calculated the percentage of differences between total scores of the HBI and P-HBI that are  $\geq 3.2$ . Fourth, to further measure agreement on the total HBI score, we examined the extent to which the clinician-based and patient-based CD activity are rated as active (HBI score  $\geq 5$ ) or as inactive (HBI score  $< 5$ )<sup>12</sup> and calculated percentage agreement. In addition, we calculated the Cohen's  $\kappa$  coefficient because, unlike percent agreement, Cohen's  $\kappa$  corrects for agreement over and above what could be expected by chance.<sup>15-17</sup> The Cohen's  $\kappa$  outcomes were interpreted as less than chance ( $< 0$ ), slight (0.01 to 0.20), fair (0.21 to 0.40), moderate (0.41 to 0.60), substantial (0.61 to 0.80), and almost perfect (0.81 to 0.99) agreement. Positive and negative predictive values were also calculated. Fifth, to assess whether having comorbid disease influenced the total sum score on the P-HBI, 2 separate analyses for patients with and without comorbidity were compared.

### Patient-Physician Agreement at Item Level

Furthermore, agreement between clinician and patient was calculated on item level. For the items with categorical response options we used Cohen's  $\kappa$  to measure agreement. One item “frequency of diarrhea” is a continuous variable. For this particular item we calculated the intraclass correlation coefficient (ICC) to measure agreement. The ICC measures agreement between 2 continuous variables, defined as the proportion of true variance compared to total variance. The interpretation of the magnitude of ICC outcomes is similar to that of Cohen's  $\kappa$ . Since 1 item, “abdominal mass,” was not included in the patient-translated HBI, agreement between clinician and patient could not be calculated for this item.

### Patient-Clinician Agreement at Total Sum and Item Level

Possibly only the single well-being item would show similar correlation as the total score of the P-HBI with the HBI. Well-being could be simpler and easier to use in clinical practice than the 11-point activity indices. Therefore, we checked whether the well-being score (1 to 10) can be used as a substitute for the P-HBI. This was done by calculating the strength of the correlation between the P-HBI item “well-being” and the total HBI scores, and

the strength of the correlation between the total scores of the P-HBI and the total HBI scores using Spearman's  $\rho$ .

### Ethical Considerations

Since no ethical approval is required for the completion of noninvasive self report questionnaires under Dutch law, the Medical Ethical Committee of the Academic Medical Centre exempted this project from formal approval.

## RESULTS

### Patient Characteristics

From April to December 2010, 205 patients at the outpatient IBD clinic were asked to complete the P-HBI. Thus, patients may have come for routine follow-up or by their own request. Twenty-four patients refused to fill out the form, because of time constraints (45.8%), lack of motivation (25%), reading constraints (16.7%), and illness (12.5%). In total, 181 patients (response rate 88.3%) with CD participated in the study and completed the P-HBI (Table 1). The median [interquartile range (IQR)] age of participants was 41 years (30 to 52 y) and the majority was female (62.4%). Moreover, CD was diagnosed at a median (IQR) age of 24 years (19 to 34 y). At the time of participation, the median (IQR) duration of CD was 12 years (6 to 21 y). In total, 19 patients (10.5%) reported to have a stoma and 127 patients (70.2%) had undergone at least 1 operation for CD in their lifetime. Less than half of the patients (47%) reported to have no comorbidity, whereas slightly more than a quarter of CD patients (27.1%) reported to have 1 comorbid disease.

### Comparison of Patient-based Versus Clinician-based HBI

To control for order effects, 74 patients (40.9%) received the P-HBI before the outpatient visit, and 107 patients (59.1%) received the P-HBI after the outpatient visit. The scores of patients who received the P-HBI before

**TABLE 1.** Sociodemographic and Clinical Patient Characteristics

	Crohn's Disease (n = 181)		
	n (%)	Median	IQR
Demographic variables			
Age (Median; range)	181	41	30-52
Sex			
Female	113 (62.4)		
Male	68 (37.6)		
Clinical characteristics			
Age at diagnosis	178*	24	19-34
Disease duration in years	178*	12	6-21
Stoma	19 (10.5)		
No. operations			
No operations	54 (29.8)		
$\geq 1$ operations	127 (70.2)		
Comorbidity			
No comorbidity	85 (47)		
1 comorbidity	49 (27.1)		
2 comorbidities	23 (12.7)		
$\geq 3$ comorbidities	22 (12.2)		

\*Numbers do not add up due to missing values.  
IQR indicates interquartile range.

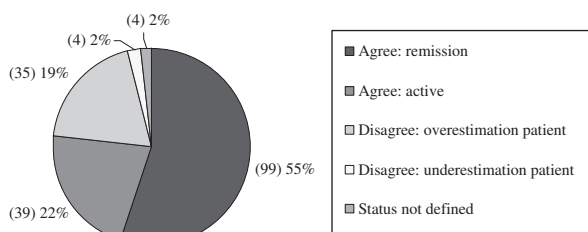
versus after the outpatient visit were not significantly different by the Mann-Whitney *U* test ( $z = 1.7, P = 0.10$ ).

**Total HBI Score**

First, Spearman’s  $\rho$  between clinician-based and patient-based total HBI score was 0.82, indicating a large correlation. Second, the median (IQR) total score of the clinician version of the HBI was 2 (0 to 4) and 4 (1 to 7) of the patient version. These total scores were significantly different according to the Mann-Whitney *U* test ( $z = 8.4, P < 0.001$ ). Repeating the comparison of the patient version with the clinician version without the abdominal mass item using the Mann-Whitney *U* test led to similar results ( $z = 8.6, P < 0.001$ ). Third, the difference between the total HBI score assessed by the patient and by the clinician was not clinically relevant (ie, <3.2 points) in 147 (82.6%) cases. Fourth, the percentage agreement between clinician and patient, both judging CD as active or as in remission, was 77%. Cohen’s  $\kappa$  yielded a score of 0.52 (moderate agreement) (Table 2 and Fig. 1). In 35 cases (19%) the P-HBI classified disease activity as active, whereas clinicians scored the same disease activity as inactive. In 4 cases (2%) the P-HBI assessed disease activity as inactive, whereas clinicians considered the same disease activity as active. Positive and negative predictive values are shown in Table 2. Fifth, the statistically significant difference between patient and clinician total HBI scores was similar for patients with and without comorbidity (data not shown).

**Item-level HBI Scores**

On the first item “well-being,” the P-HBI correlated highly ( $r_s = 0.65$ ) with the clinician-based HBI and yielded a moderate agreement ( $\kappa = 0.52$ ). Compared to the clinician-based assessment, the P-HBI scored higher on well-being in 20 cases, whereas patients scored lower on well-being in 28 cases (Table 3). The second item “abdominal pain” showed a high correlation ( $r_s = 0.53$ ) and a slight agreement ( $\kappa = 0.14$ ) between clinician and patient assessment. In 109 cases patients scored “abdominal pain” on the P-HBI higher than clinicians, whereas abdominal pain was never scored lower by patients (Table 4). The third item “frequency of diarrhea” yielded a substantial agreement (ICC = 0.80; 95% CI, 0.74-0.85) between the 2 questionnaires. Agreement between the patient-based and clinician-based HBI on frequent extraintestinal manifestations of CD varied from less than chance to a perfect agreement (Table 5). Erythema nodosum ( $\kappa = 1$ ) yielded a perfect agreement, arthralgia ( $\kappa = 0.56$ ), active fistula ( $\kappa = 0.49$ ), and uveitis ( $\kappa = 0.49$ ) a moderate agreement, aphthous ulcer ( $\kappa = 0.35$ ) a fair



**FIGURE 1.** Proportion of (dis)agreement regarding disease activity.

agreement, fissure ( $\kappa = 0.03$ ) a slight agreement, pyoderma ( $\kappa = 0$ ) no agreement, and abscess ( $\kappa = -0.02$ ) an agreement below chance. We repeated all analyses without the patients with a pouch or a stoma. This led to similar results (data not shown).

**Patient-Clinician Agreement at Total Sum and Item Level**

The item “well-being” of the P-HBI correlated lower with the total scores of the HBI ( $r_s = -0.55$ ) than the total scores of the P-HBI ( $r_s = 0.82$ ) did.

**DISCUSSION**

This study compared a patient-based P-HBI as assessed by CD patients with the HBI as assessed by the treating gastroenterologist. If high agreement is found, the P-HBI can easily be applied in daily clinical practice, may facilitate early detection of relapse, would be helpful for large scale future research, and last but not least reinforce patient empowerment.

**Agreement Between HBI and P-HBI**

The P-HBI appears to highly correlate with clinician-derived scores on the total HBI score. Although patients had significantly higher HBI total scores than clinicians, only a small percentage (17%) represented a clinically significant difference. Agreement between clinicians and patients in assessing CD as active or as in remission was moderate to large. Moreover, the positive predictive value of P-HBI is noteworthy. When patients judge CD to be in remission, in 96% of the cases clinicians will also judge the disease as in remission. In contrast, the negative predictive value of CD is lower. When patients assessed their disease as active, only about half of the clinicians agreed with the patient’s judgment.

**TABLE 2.** Clinician-Patient Association and Agreement on HBI Judged as in Remission or as Active

		Clinician Assessment		
		Remission (< 5)	Active (≥ 5)	
Patient assessment	Remission (< 5) (= positive)	99	4	Positive predictive value = 0.96 Negative predictive value = 0.53
	Active (≥ 5) (= negative)	35	39	
Agreement	77%	$r_s = 0.56$	$\kappa = 0.52$	

\*Due to missing values.  
Data are presented as frequencies unless stated otherwise.  
HBI indicates Harvey Bradshaw Index;  $r_s$  = Spearman rank correlation;  $\kappa$  = Kappa.

**TABLE 3.** Clinician-Patient Association and Agreement on HBI Item “Well-being” on a 10-Point Scale

n = 180*		Clinician Assessment				
		< 4	4	5	6	≥ 7
Patient assessment	< 4	2	3	1	0	1
	4	0	1	2	2	1
	5	0	0	5	8	2
	6	0	0	0	29	8
	≥ 7	0	2	1	17	95
Agreement	73%	$r_s = 0.65$			$\kappa = 0.52$	

\*Due to missing values.

Data are presented as frequencies unless stated otherwise.

HBI indicates Harvey Bradshaw Index;  $r_s$  = Spearman rank correlation;  $\kappa$  = Kappa.

**Agreement on Item-level HBI Scores**

In addition, we looked at differences between patients and clinicians at item level. The question on frequency of diarrhea correlated highly between the P-HBI and HBI, indicating that the number of stools per day can be easily assessed by the patient. Considering the questions concerning extraintestinal manifestations, erythema nodosum, uveitis, and pyoderma were not very common among our patients, no firm conclusion can be drawn for these items. Other items did indicate a clear trend.

When the P-HBI and HBI differ, patients report a higher disease activity than the gastroenterologists: more abdominal pain, more extraintestinal manifestations (ie, arthralgia, active fistula, aphthous ulcer, anal fissure, and abscess), and a lower level of well-being.

Previous studies on agreement between patients and health care providers performed in a wide range of patient populations and in various settings, report mixed findings.<sup>18-21</sup> A review evaluating health care providers' appraisal of quality of life of patients with chronic diseases reported that providers overestimated patients' symptomatology.<sup>18</sup> As most studies included a limited number of patients these results must be interpreted with caution. Another study evaluating the extent of agreement between patients and health care providers concluded that providers underestimated symptom intensities (eg, pain, fatigue, depression, constipation, and diarrhea) of patients.<sup>19</sup> Again, care must be taken when interpreting these results, as only cancer patients were included in this study. A recent survey on quality of life in IBD including 2424 patients, the largest IBD study to date, also found that patients reported more symptoms and a larger impact of the disease on their lives than clinicians did.<sup>19</sup> Indeed, our findings are in line with the latter 2 studies.

The discrepancies found in our study might either reflect an overestimation of the patients while filling out the questionnaire or an underestimation of the clinicians during consultation. First, clinicians and patients have a different standard of comparison, as clinicians observe many IBD patients, whereas patients lack such a comparative perspective.<sup>18</sup> Second, some patients may be reluctant to answer the questions accurately when asked by a clinician. For instance, during the consultation patients might report less abdominal pain to avoid hospitalization or physical examination, such as an invasive endoscopy. In addition, patients might give socially desirable answers. This might be particularly relevant for the item on well-being. Patients may assume that the task of the clinician is merely to address physical symptoms and not symptoms related to their well-being. Alternatively, clinicians may not thoroughly probe patients about their well-being and only base their judgment on physical aspects.<sup>22</sup> Finally, discrepancies in agreement between patients and clinicians may be attributed to the level of objectivity of symptoms. Abdominal pain and well-being are more subjective variables than diarrhea.

**Patient-Clinician Agreement at Total Sum and Item Level**

Since it was assumed that the patient's well-being score might be easier to use in clinical practice, we checked whether it can replace the total score of the P-HBI. However, results show that although the well-being item moderately correlates with the HBI total scores, the total scores of the P-HBI have a much larger correlation with the HBI. Therefore we are in favor of using the total scores of the P-HBI.

**TABLE 4.** Clinician-Patient Association and Agreement on HBI Item “Abdominal Pain” on a 10-Point Scale

n = 179*		Clinician Assessment			
		No	1-3	4-7	8-10
Patient assessment	No	49	0	0	0
	1-3	52	13	0	0
	4-7	26	19	8	0
	8-10	3	5	4	0
Agreement	39%	$r_s = 0.53$		$\kappa = 0.14$	

\*Due to missing values.

Data are presented as frequencies unless stated otherwise.

HBI indicates Harvey Bradshaw Index;  $r_s$  = Spearman rank correlation;  $\kappa$  = Kappa.

**TABLE 5.** Clinician-Patient Agreement on HBI Items “Extraintestinal Manifestations”

Patient/Clinician Association and Agreement	N	% Total Agreement	Extraintestinal Manifestations		Patient Lower Scores (n)	Patient Higher Scores (n)	<i>r<sub>s</sub></i>	$\kappa$
			“No” Agreement (n)	“Yes” Agreement (n)				
Erythema nodosum	181	100	180	1	0	0	1	1
Arthralgia	179*	81	108	37	5	29	0.59	0.56
Active fistula	181	93	161	7	0	13	0.57	0.49
Uveitis	181	99	178	1	1	1	0.49	0.49
Aphthous ulcer	181	96	172	2	1	6	0.39	0.35
Fissure	180*	82	147	1	2	30	0.06	0.03
Pyoderma gangrenosum	181	99	180	0	0	1	0	0
Abscess	180*	95	171	0	2	7	-0.02	-0.02

\*Due to missing values.

HBI indicates Harvey Bradshaw Index; *r<sub>s</sub>* = Spearman rank correlation;  $\kappa$  = Kappa.

### Clinical Implications and Recommendations for Further Research

The current study found that when CD is in remission the clinician’s assessment can be replaced by the patient’s assessment (P-HBI). However, when the disease is active, only about half of the clinicians agreed with the patient’s judgment. Therefore, when CD is active results obtained from the P-HBI and HBI must be approached with caution. An incorrect interpretation of symptoms may lead to inappropriate treatment. Solely consulting the HBI or P-HBI may result in undertreatment or overtreatment and unnecessary symptom burden. In order to investigate which of the 2 questionnaires is more valid, future studies should compare the P-HBI and HBI with prospective biological markers (eg, endoscopy or cross-sectional imaging) of disease activity in the entire sample. Furthermore, future research is needed to investigate the influence of patient and clinical characteristics on the agreement between patients and physicians with respect to disease activity.

### Limitations of the Study

Our study bears a few limitations. First, the comparison between the patient-based and clinician-based assessments of the HBI was performed in a tertiary referral center, where severely ill CD patients are likely to be over-represented. Therefore, our results should not be generalized to patients with more mild forms of CD attending peripheral hospitals or receiving primary care. Second, as only 3 clinicians participated in our study, generalization to the larger clinician population might be limited. Third, most CD extraintestinal manifestations were not very common among our cohort of patients. Hence, future research should replicate our study with a larger and more heterogeneous sample of clinicians and patients including those patients with mild and severe CD and with and without CD extraintestinal manifestations and register their disease history (eg, medications used and Montreal Classification). Fourth, contrary to the clinicians’ version of the HBI our P-HBI did not include the item about abdominal mass, whereas our cut-off scores<sup>11</sup> were based on the original HBI that does include this item. Fortunately, the item “abdominal mass” was scored by the clinician in only 2.7% (n = 5) of the current patient group. We repeated our analyses with and without the abdominal mass item for clinicians, which led to similar results. Therefore, it is

unlikely that the absence of the item on abdominal mass on the P-HBI had a significant impact on our findings.

### Strengths

This study also has several strengths. First and foremost, it is the first study that has developed a patient-based questionnaire of the HBI, and has compared this questionnaire with a clinician-based assessment. The P-HBI can easily be completed by patients without the presence of a clinician allowing its use as a web-based questionnaire. Furthermore, to the best of our knowledge, this is the first study evaluating the extent of patient and health care provider agreement in patients with CD, whereas 2 previous studies evaluated the extent of agreement in patients with ulcerative colitis.<sup>23,24</sup> Third, our P-HBI incorporates clear instructions and descriptions to facilitate a standardized and undemanding patient self-administration of disease activity. Moreover, the scoring system is simple and equivalent to the original clinician-based HBI. A fourth strength is that the patient-based and clinician-based assessment of the HBI took place on the same day within a time span of an hour. Therefore, it is to be expected that disease activity did not change and thus that the patient-based and clinician-based assessment captured the same CD activity. A fifth strength is that clinicians were blinded for the patients’ responses. Finally, to control for order effects patients received the P-HBI at random before or after the consultation. Whereas more patients completed the P-HBI after the consultations due to clinicians’ time constraints, the potential order effect was not found to influence our results.

### CONCLUSIONS

This study represents a first step in the development of a patient-based translation of the HBI. The current version of this patient-based HBI has a high clinical agreement (96%) with the clinician-based HBI when CD is in remission. The P-HBI can be used as a complementary assessment of disease activity and has additional value by enquiring into different symptoms of disease activity and extra intestinal manifestations. It might help the clinician to preselect patients who have to be invited for a clinical consult. If patients score low on the P-HBI there is a high probability (96%) that disease activity is also low according to the clinicians. Therefore, for these patients a clinical consult might be postponed. However, further research is necessary when the P-HBI indicates an active disease to

avoid placing patients at risk for undertreatment or over-treatment. These patients with high scores need to be seen by their clinician using the original HBI. If CD activity can be measured without the patient having to visit the gastroenterologist or to undergo physical examination, this might have a positive impact on their quality of life. Using a self-report assessment is less demanding for patients in routine clinical care and will facilitate clinical research. Moreover, patient-based assessment of CD activity after having improved the P-HBI for high scores ( $\geq 5$ ) would increase the chance of early detection of relapse. Furthermore, the P-HBI can potentially reduce health care costs, by reducing visits to the gastroenterologist and requiring fewer physical examinations or laboratory tests.

Further research is essential to validate the P-HBI in a variety of patient populations, including patients with a recent onset of CD, before this can be used in daily clinical practice. Moreover, future research should further explore the reliability, validity, and responsiveness of the P-HBI by comparing it to biological markers of disease activity.

### ACKNOWLEDGMENTS

The authors thank the gastroenterologists and specially Prof. J.F.W.M. Bartelsman, and M. Lowenberg, PhD, of the outpatient clinic of the Academic Medical Center, for their collaboration and recruitment of patients for this study. The authors acknowledge the contribution of the research assistant K. Sitnikova, MSc, for the review of earlier drafts of the article.

### REFERENCES

- Baumgart DC, Bernstein CN, Abbas Z, et al. IBD around the world: comparing the epidemiology, diagnosis, and treatment: Proceedings of the World Digestive Health Day 2010—Inflammatory Bowel Disease Task Force Meeting. *Inflamm Bowel Dis*. 2011;17:639–644.
- Loftus EV. Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. *Gastroenterology*. 2004;126:1504–1517.
- Loftus CG, Loftus EV Jr, Harmsen WS, et al. Update on the incidence and prevalence of Crohn's disease and ulcerative colitis in Olmsted County, Minnesota, 1940–2000. *Inflamm Bowel Dis*. 2007;13:254–261.
- Graff LA, Walker JR, Lix L, et al. The relationship of inflammatory bowel disease type and activity to psychological functioning and quality of life. *Clin Gastroenterol Hepatol*. 2006;4:1491–1501.
- Von Wietersheim J, Kessler H. Psychotherapy with chronic inflammatory bowel disease patients: a review. *Inflamm Bowel Dis*. 2006;12:1175–1184.
- Husain A, Triadafilopoulos G. Communicating with patients with inflammatory bowel disease. *Inflamm Bowel Dis*. 2004;10:444–450.
- Yoshida EM. The Crohn's Disease Activity Index, its derivatives and the Inflammatory Bowel Disease Questionnaire: a review of instruments to assess Crohn's disease. *Can J Gastroenterol*. 1999;13:65–73.
- Sands BE, Ooi CJ. A survey of methodological variation in the Crohn's disease activity index. *Inflamm Bowel Dis*. 2005;11:133–138.
- Jørgensen LGM, Fredholm L, Petersen PH, et al. How accurate are clinical activity indices for scoring of disease activity in inflammatory bowel disease (IBD)? *Clin Chem Lab Med*. 2005;43:403–411.
- Best WR. Predicting Crohn's Disease Activity Index from the Harvey-Bradshaw Index. *Inflamm Bowel Dis*. 2006;12:304–310.
- Harvey RF, Bradshaw JM. A simple index of Crohn's-Disease activity. *The Lancet*. 1980;1:514.
- Vermeire S, Schreiber S, Sandborn WJ, et al. Correlation between the Crohn's Disease Activity and Harvey-Bradshaw indices in assessing Crohn's Disease severity. *Clin Gastroenterol Hepatol*. 2010;8:357–363.
- Sandler RS, Jordan MC, Kupper LL. Development of a Crohn's index for survey research. *J Clin Epidemiol*. 1988;41:451–458.
- Thia K, Faubion WA, Loftus EV, et al. Short CDAI: development and validation of a shortened and simplified Crohn's Disease Activity Index. *Inflamm Bowel Dis*. 2011;17:105–111.
- Boushka WM, Marinez YN, Prihoda TJ, et al. A computer program for calculating kappa: application to interexaminer agreement in periodontal research. *Comput Methods Programs Biomed*. 1990;33:35–41.
- Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85:257–268.
- Rigby AS. Statistical methods in epidemiology. V. Towards an understanding of the kappa coefficient. *Disabil Rehabil*. 2000;22:339–344.
- Sneeuw KCA, Sprangers MAG, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. *J Clin Epidemiol*. 2002;55:1130–1143.
- Laugsand EA, Sprangers MAG, Bjordal K, et al. Health care underestimate symptom intensities of cancer patients: a multi center European study. *Health Qual Life Outcomes*. 2010;8:104–117.
- Lesage AC, Hagege H, Tucut G, et al. Results of a national survey on quality of life in inflammatory bowel diseases. *Clinics Res Hepatol Gastroenterol*. 2011;35:117–124.
- Jones JM, McPherson CJ, Zimmermann C, et al. Assessing agreement between terminally ill cancer patients' reports of their quality of life and family caregiver and palliative care physician proxy ratings. *J Pain Symptom Manage*. 2011;42:354–365.
- Yen JC, Abrahamowicz M, Dobkin PL, et al. Determinants of discordance between patients and physicians in their assessment of lupus disease activity. *J Rheumatol*. 2003;30:1967–1976.
- Lee JJ, Colman RJ, Mitchell DP, et al. Agreement between patient- and physician-completed pediatric ulcerative colitis activity index scores. *J Pediatr Gastroenterol Nutr*. 2011;52:708–713.
- Turner D, Griffiths AM, Mack D, et al. Assessing disease activity in ulcerative colitis: patients or their physicians? *Inflamm Bowel Dis*. 2010;16:651–656.