

A prospective study comparing patient-reported outcomes in Crohn's disease

Daniël R. Hoekman^{a,b}, Mark Löwenberg^a, Gijs R. van den Brink^a, Cyriel Y. Ponsioen^a, Marc A. Benninga^b and Geert R. D'Haens^a

Background Patient reported outcomes are important in Crohn's disease. In this prospective cohort, we investigated the performance of the Bristol Stool Form Scale (BSFS) and a visual analog scale (VAS) for abdominal pain as outcome measures in Crohn's disease.

Methods Patients with active Crohn's disease starting glucocorticoids or anti-tumor necrosis factor were included. Before treatment and 10 weeks later we collected: clinical activity [Harvey Bradshaw Index (HBI) and Crohn's-Disease-Activity-Index (CDAI)], serum C-reactive protein (CRP) and fecal calprotectin, and BSFS (1–7) and a 100-mm VAS based on a 7-day diary. Clinical response was defined as a reduction by at least 3 and at least 100 of HBI and CDAI, respectively. Fecal calprotectin-response and CRP-response were defined as reduction of at least 50%.

Results Thirty-eight patients completed follow-up. At baseline, BSFS-parameters correlated more strongly with clinical activity (range: r_s : 0.31–0.74) than with CRP (r_s : –0.01 to 0.16) and fecal calprotectin (r_s : 0.14–0.26). VAS scores correlated very weakly to moderately with clinical activity (r_s : 0.18–0.45), and weakly to moderately with CRP (r_s : 0.24–0.34) and fecal calprotectin (r_s : 0.35–0.43). Changes in VAS scores correlated moderately to strongly (r_s : 0.55–0.71) with changes in clinical activity, and weakly with changes in CRP and fecal calprotectin (r_s : 0.21–0.35). Changes in BSFS parameters correlated weakly to moderately (r_s : 0.23–0.53) with changes in clinical activity, and very weakly to weakly (r_s : 0.01–0.35) with changes in CRP and fecal calprotectin. Responsiveness of VAS and BSFS was moderate to high (Guyatt's statistic 0.41–2.17) and highly dependent on the definition of response.

Conclusions The BSFS and a VAS appear to be responsive with moderate-to-strong construct validity to monitor patients with Crohn's disease. *Eur J Gastroenterol Hepatol* 32: 38–44

Copyright © 2019 The Author(s). Published by Wolters Kluwer Health, Inc.

Introduction

Traditionally, the efficacy of therapeutic interventions in clinical trials in Crohn's disease was assessed using the Crohn's Disease Activity Index (CDAI), an index consisting of laboratory tests, physical examination and patient-reported symptoms based on a 7-day diary [1]. Although the CDAI correlates well with the physician's overall global assessment of disease activity [1], it is not disease specific and it correlates very poorly with endoscopic severity [2], an important predictor of long-term outcome [3]. Consequently, it has been suggested that

mucosal healing in combination with patient-reported outcomes (PROs) should be a primary outcome measure in clinical trials [4,5].

Because the CDAI includes components that are not reported by patients (abdominal mass assessed by a physician, body weight, hematocrit), it does not qualify as a genuine PRO. Furthermore, the symptom-related items in the CDAI were selected based on the gastroenterologist's point-of-view [1], rather than on what is important for the patient, and the items' weights were arbitrarily chosen rather than based on rigorous biometric techniques.

PROs can be defined as any report of the status of a patient's health condition that comes directly from the patient without interpretation of the patient's response by a clinician or anyone else [6], and are increasingly recommended to be used as outcome measures in clinical trials and daily practice. Recently, the United States Food and Drug Administration (FDA) proposed guidance for the development of PROs [6]. According to this guidance, the CDAI is no longer acceptable as a primary outcome in clinical trials [6].

However, the development of PROs according to the FDA guidance is a lengthy process involving patient concept elicitation interviews, expert interviews, item generation, content validity testing, patient cognitive interviews and quantitative validation [7]. Currently, no PRO instrument is available for Crohn's disease which is developed in accordance with the FDA guidance.

In clinical practice and in trials, the Bristol Stool Form Scale (BSFS) and a visual analogue scale (VAS) are frequently

European Journal of Gastroenterology & Hepatology 2020, 32:38–44

Keywords: Bristol stool form scale, Bristol stool chart, Crohn's disease, patient-reported outcomes, visual analog scale

Departments of ^aGastroenterology and Hepatology and ^bPediatric Gastroenterology and Nutrition, Amsterdam University Medical Center, Amsterdam, The Netherlands

Correspondence to Geert R. D'Haens, MD, PhD, Amsterdam University Medical Center, PO Box 22700, 1100 DD Amsterdam, The Netherlands
Tel: +3120 5661768; fax: +3120 6917033; e-mail: g.dhaens@amc.uva.nl

Received 25 April 2019 **Accepted** 7 August 2019

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.eurojgh.com.

This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CC-BY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

used to evaluate bowel habits and symptoms of abdominal pain and discomfort. To our knowledge, these instruments have never been evaluated as outcome measures in Crohn's disease, although they are truly 'patient reported'. We hypothesized that the BSFS and VAS scores for abdominal pain correlate well with clinical and biochemical disease activity, and that these measures could be used as outcome measures in Crohn's disease. In this prospective study, we investigated whether the BSFS and a VAS for abdominal pain can be used as outcome measures in Crohn's disease. Therefore, we examined the criterion validity and responsiveness of these instruments in patients with Crohn's disease who start tumor necrosis factor (TNF) inhibitors or corticosteroid treatment (treatments of known efficacy) for active disease.

Methods

Population

We performed a single center, prospective, observational study at the Academic Medical Center in Amsterdam, The Netherlands, between October 2013 and February 2016. Eligible patients were 18 years of age or older with an endoscopically and histologically confirmed diagnosis of Crohn's disease, scheduled to start oral corticosteroids or TNF inhibitors for active disease based on gastrointestinal symptoms in combination with biochemical evidence of inflammation [high-sensitivity serum C-reactive protein (CRP) > 5 mg/L and/or fecal calprotectin > 250 µg/g]. Patients with an ostomy or a history of colectomy were excluded. Oral prednisolone was administered at a daily dose of 40 mg for 3 weeks (irrespective of body weight), followed by a gradual tapering by 5 mg every week in the case of clinical response. Budesonide was administered orally at a daily dose of 9 mg for 8 weeks, followed by a 2-week period at a daily dose of 6 mg. Infliximab was administered intravenously at a dose of 5 mg per kg body-weight at week 0, 2 and 6. Adalimumab was administered subcutaneously at 160 mg at week 0, 80 mg at week 2, followed by 40 mg every other week.

Outcome measures

Comprehensive measurement of disease activity and PROs was performed prior to treatment and 10 weeks later. Study outcomes that were studied were divided in three categories: (1) the 'conventional' disease activity scores: CDAI and Harvey Bradshaw Index (HBI); (2) biochemical markers of disease activity: high sensitivity serum CRP and fecal calprotectin (Bühlmann ELISA, Schönenbuch, Switzerland); and (3) PROs: BSFS, a visual analogue abdominal pain scale, PRO2 and PRO3 (see below).

During the week prior to treatment and 10 weeks after the initiation of treatment, participants were asked to complete a diary (provided in supplements, Supplemental digital content 1, <http://links.lww.com/EJGH/A464>) for seven consecutive days for daily recording of the number of loose/watery stools, abdominal pain, general well-being (the three components of the CDAI), abdominal pain on a 100 mm VAS, and for the daily scoring of the appearance of the feces: patients were asked to score which types of stools they have had on the BSFS (1–7). HBI scores were recorded on the day of the visit, reflecting symptoms on the day prior to the visit. The PRO measures PRO2 and PRO3

were calculated from CDAI diaries, as described previously. For each diary component of the CDAI, the daily average score over 7 days was weighted using the original CDAI multiplication factors. Subsequently, the scores for all items (number of loose/watery stools, abdominal pain, ± general well-being) were summed to create PRO2 and PRO3, respectively [8].

Clinical response was based on HBI and CDAI scores. HBI response was defined as a reduction in HBI by ≥3 points [9]. CDAI response was defined as a reduction in CDAI ≥100 points, as frequently used in clinical trials [10]. Clinical remission was defined as a CDAI score of <150 and/or an HBI score of ≤4. Biochemical response was defined as a reduction of ≥50% compared with baseline values of fecal calprotectin (fecal calprotectin response) and/or CRP (CRP response), as used previously [11]. Diarrhea on the BSFS was defined as a score of ≥6. Severe diarrhea was defined as a BSFS score of 7. Constipation was defined as a BSFS score of ≤2 [12].

Sample size

The primary aim of the study was to evaluate the correlation between changes in BSFS and VAS score with changes in clinical and biochemical outcomes as a measure of criterion validity. A sample of 36 patients is generally deemed sufficient to detect an anticipated correlation coefficient of 0.45 different from 0.00 with 80% power at the 0.05 (two-sided) level of significance.

Statistical analysis

At baseline, Spearman's rank correlation of the following outcomes with clinical scores (HBI, CDAI, PRO2 and PRO3) and biochemical outcomes (CRP, fecal calprotectin) were studied: mean VAS pain score over 1 week, VAS score at day 1 of the diary (D1), proportion of days without diarrhea based on the BSFS, proportion of days without severe diarrhea based on the BSFS, the highest BSFS score at D1 and the highest BSFS score during the whole week. Furthermore, Spearman's rank correlation of the following outcomes was evaluated: change in mean VAS score over 1 week, VAS score at D1, proportion of days without diarrhea based on the BSFS, proportion of days without severe diarrhea based on the BSFS, highest BSFS score at D1 and highest BSFS score over 1 week. Mean BSFS scores were not analyzed, because they can be misleading [e.g., the mean score of constipation (BSFS = 1) and liquid stools (BSFS = 7) are normal stool (BSFS = 4)]. Correlation coefficients were interpreted as follows: 0–0.2 very weak; 0.2–0.4 weak; 0.4–0.6 moderate; 0.6–0.8 strong; 0.8–1.0 very strong [13]. Guyatt's responsiveness statistic was calculated as the mean change in responders divided by the SD of the change in nonresponders. Values of 0.2–0.5, 0.5–0.8 and ≥0.8 were considered to represent low, moderate and high responsiveness, respectively [14].

Missing data

In the case of missing days in the diary, data from the completed days were analyzed. If more than 4 days of a diary were missing, patients were excluded from the analysis. In the case of missing CRP and/or fecal calprotectin levels, patients were excluded from the corresponding analyses

only. In the case of missing hematocrit levels, the CDAI score was calculated based on the last observation carried forward principle using the last recorded hematocrit level.

Ethical considerations

The study protocol was reviewed by the local ethics committee, who deemed that, in line with the Dutch Medical Research Involving Human Subjects Act (Wet Medisch-Wetenschappelijk Onderzoek met Mensen), full ethical review was not required.

Results

Study population

Fifty-one patients participated in the study, of whom 38 completed the study and were included in these analyses. Baseline characteristics are provided in Table 1. Reasons for not completing the study were noncompliance with diaries ($n=7$), withdrawal of consent ($n=3$), loss to follow-up ($n=2$) and a revised diagnosis (no Crohn's disease; $n=1$). The median CDAI score at baseline was 173 points interquartile range [(IQR) 125–225]. Median CRP and fecal calprotectin levels were 4.9 mg/L (IQR 2.4–15.2) and 825 $\mu\text{g/g}$ (291–1800), respectively.

Treatment efficacy

Corticosteroids were started in 24 (63%) patients [7 (18%) prednisolone, 17 (45%) budesonide], and TNF inhibitors in 14 (37%) patients [6 (16%) infliximab, 8 (21%) adalimumab]. Clinical response at week 10 based on HBI and CDAI was observed in 14 (37%) and 8 (21%) patients, respectively, while CRP and fecal calprotectin response was seen in 24 (63%) and 16 (47%) patients, respectively. The proportion of patients in remission at week 10 was 63%, based on either HBI or CDAI.

Table 1. Baseline characteristics

Patients (n)	38
Age, years (mean, SD)	40.7 (14.6)
Sex, male (n , %)	12 (31%)
Age at diagnosis (n , %)	
A1 (<17 years)	4 (11%)
A2 (17–40 years)	28 (74%)
A3 (>40 years)	6 (16%)
Disease location (n , %)	
L1 (ileal)	5 (13%)
L2 (colonic)	6 (16%)
L3 (ileocolonic)	27 (71%)
L4 (upper gastrointestinal tract)	3 (8%)
Disease behavior (n , %)	
B1 (nonstricturing nonpenetrating)	25 (66%)
B2 (stricturing)	10 (26%)
B3 (penetrating)	3 (8%)
P (perianal disease)	8 (21%)
Prior bowel resection	9 (24%)
CRP (median, IQR) at baseline	4.9 (2.4–15.2) mg/L
Fecal calprotectin (median, IQR) at baseline	825 (291–1800) $\mu\text{g/g}$
CDAI (median, IQR) at baseline	173 (125–225)
HBI (median, IQR) at baseline	6 (4–8)
PRO2 (median, IQR) at baseline	11 (8–17)
PRO3 (median, IQR) at baseline	27 (22–36)
Use of immunomodulators (n , %)	16 (42%)
AZA/6-MP/6-TG	13 (34%)
MTX	3 (8%)

BSFS, Bristol stool form scale; CDAI, Crohn's disease activity index; CRP, C-reactive protein; IQR, interquartile range; HBI, Harvey Bradshaw index; MTX, methotrexate; PRO, patient-reported outcome; VAS, visual analog scale.

Bristol Stool Form Scale

Descriptive statistics

Maximum BSFS scores at baseline and after treatment are provided in Fig. 1. An overview of all reported BSFS scores is provided in Supplementary Fig. 1, Supplemental digital content 2, <http://links.lww.com/EJGH/A465>. At baseline, the majority of patients had diarrhea and severe diarrhea at day 1 ($n=29$, 76%; $n=19$, 50%, respectively), or at least one day during the week ($n=34$, 89%; $n=26$, 68%, respectively). After treatment, 28 (74%) patients still suffered from diarrhea at least one day during the week, and 14 (37%) patients had severe diarrhea at least once during the week.

Prior to treatment, 6 (16%) patients had constipation-type stools at day 1, and 10 (26%) patients had constipation-type stools at least one day during the week. This decreased to 2 (5%) and 8 (21%) after treatment, respectively.

Relationship with disease activity

The correlation of BSFS scores at baseline with clinical disease activity index scores, PRO2, PRO3, fecal calprotectin and CRP is provided in Table 2. The highest BSFS score correlated weakly to moderately with clinical disease activity (r_s : 0.31–0.54) and weakly to very weakly with CRP (r_s : –0.01 to 0.16) and fecal calprotectin (r_s : 0.14–0.27). The proportion of days with diarrhea or severe diarrhea correlated best with CDAI (r_s : 0.42–0.50) and HBI (r_s : 0.51–0.69). BSFS outcomes correlated more strongly with PRO2 score (r_s : 0.54–0.74) compared to PRO3 score (r_s : 0.29–0.53).

Criterion validity and responsiveness to change

The correlation between changes in BSFS scores and changes in clinical disease activity index scores, PRO2, PRO3, fecal calprotectin and CRP after treatment compared with baseline is provided in Table 3 and Fig. 2a–f. Changes in BSFS outcomes correlated weakly to moderately to changes in clinical disease activity indices (r_s : 0.23–0.53), and weakly to very weakly with changes in fecal calprotectin (r_s : 0.10–0.35) and CRP (r_s : 0.01–0.15). At week 10, similar to baseline, changes in BSFS outcomes correlated more strongly with changes in PRO2 score (r_s : 0.47–0.63) compared with changes in PRO3 score (r_s : 0.27–0.56).

Guyatt's responsiveness statistics of changes in proportion of days with diarrhea with respect to response based on CDAI and HBI were respectively 1.82 and 1.48, indicating high responsiveness and 0.68 and 0.70 for response based on fecal calprotectin and CRP, respectively, indicating moderate responsiveness. Guyatt's statistics of changes in proportion of days with severe diarrhea were 1.57, 1.12, 0.41 and 0.28, respectively. Guyatt's statistics of highest BSFS score during the whole week were 1.10, 1.40, 0.60 and 0.57, respectively. Guyatt's statistics of highest BSFS score at day 1 were 1.39, 1.36, 0.61 and 0.69, respectively.

Visual analog scale for abdominal pain

Descriptive statistics

At baseline, the mean VAS score at day 1 was 36 mm (SD: 23), and mean average VAS score during the week was 34 mm (SD: 21). After treatment, the mean VAS score on

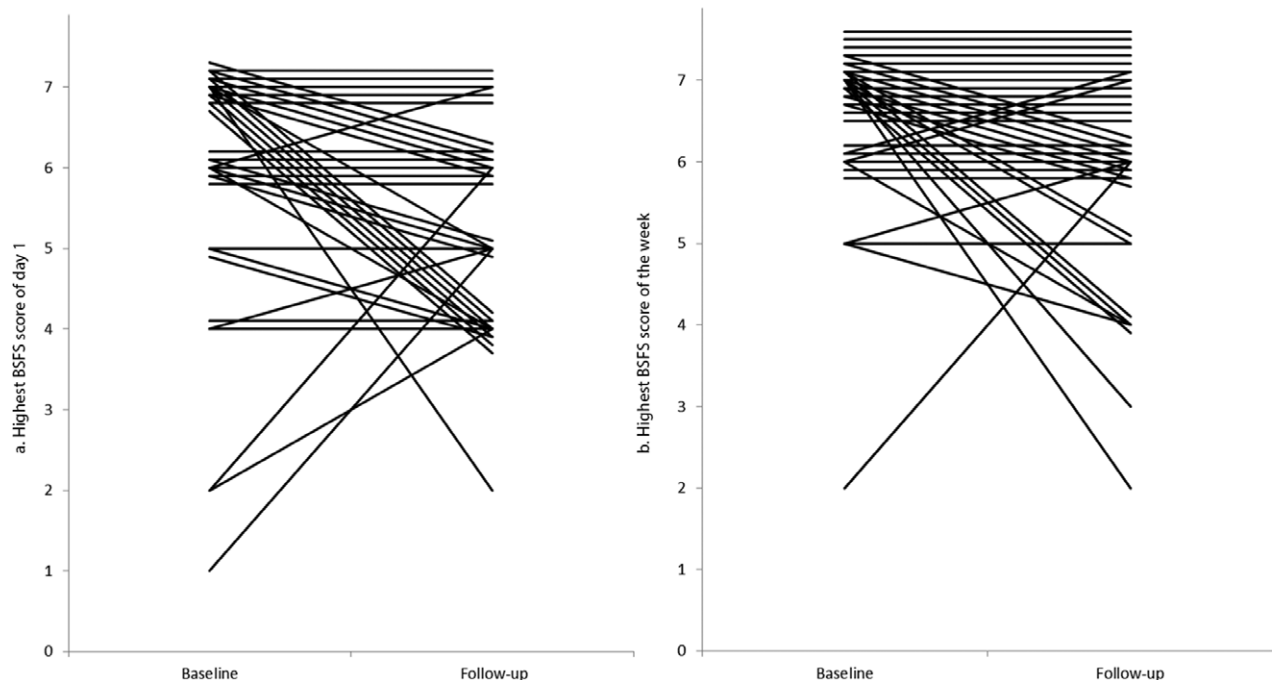


Fig. 1. Highest BSFS score of (a) day 1 and (b) week 1 before and after treatment. Lines represent individual patients. All values represent absolute numbers. Lines have been moved vertically to allow for the visualization of overlapping data. BSFS, Bristol stool form scale.

Table 2. Spearman's rank correlation coefficients of Bristol stool form scale score, visual analog scale score, Crohn's disease activity index, Harvey Bradshaw index, patient-reported outcome 2, patient-reported outcome 3, fecal calprotectin and C-reactive protein at baseline

	% Days with severe diarrhea	Highest BSFS score (1 day)	Highest BSFS score (1 week)	Mean VAS score (1 week)	VAS score (day 1)	CDAI	HBI	PRO2	PRO3	Fecal calprotectin	CRP
% Days with diarrhea	0.71***	0.78***	0.62***	0.02	-0.07	0.50**	0.69***	0.74***	0.53***	0.25	-0.02
% Days with severe diarrhea		0.83***	0.81***	0.14	0.08	0.42**	0.51***	0.59***	0.38*	0.18	-0.09
Highest BSFS score (D1)			0.71**	0.10	0.08	0.31	0.54***	0.54***	0.34*	0.27	-0.01
Highest BSFS score (1 week)				0.16	0.05	0.33*	0.37*	0.55***	0.29	0.14	-0.16
Mean VAS score (1 week)					0.86***	0.35*	0.18	0.33*	0.45**	0.34	0.43**
VAS score (D1)						0.27	0.20	0.19	0.31	0.24	0.35*
CDAI							0.79***	0.82***	0.95***	0.27	0.28
HBI								0.80***	0.78***	0.16	0.16
PRO2									0.87***	0.30	0.12
PRO3										0.34*	0.28
Fecal calprotectin											0.54*

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

BSFS, Bristol stool form scale; CDAI, Crohn's disease activity index; CRP, C-reactive protein; HBI, Harvey Bradshaw index; PRO, patient-reported outcome; VAS, visual analog scale.

Table 3. Spearman's rank correlation coefficients of changes in Bristol stool form scale score, visual analog scale score, Crohn's disease activity index, Harvey Bradshaw index, patient-reported outcome 2, patient-reported outcome 3, fecal calprotectin and C-reactive protein between baseline and after treatment

	Δ % days with severe diarrhea	Δ Highest BSFS score (1 day)	Δ Highest BSFS score (1 week)	Δ Mean VAS score (1 week)	Δ VAS score (day 1)	Δ CDAI	Δ HBI	Δ PRO2	Δ PRO3	Δ Fecal calprotectin	Δ CRP
Δ % days with diarrhea	0.65***	0.69***	0.69***	0.39*	0.28	0.50**	0.43**	0.63***	0.56***	0.35*	0.15
Δ % days with severe diarrhea		0.63***	0.56***	0.22	0.23	0.29	0.38*	0.52**	0.36*	0.10	-0.11
Δ Highest BSFS score (1 day)			0.70***	0.24	0.29	0.34*	0.53**	0.53**	0.41*	0.20	0.11
Δ Highest BSFS score (1 week)				0.11	0.14	0.23	0.41*	0.47**	0.27	0.12	0.01
Δ Mean VAS score (1 week)					0.77***	0.65***	0.58***	0.71***	0.71***	0.23	0.35*
Δ VAS score (day 1)						0.55***	0.71***	0.59***	0.57***	0.21	0.21
Δ CDAI							0.75***	0.82***	0.95***	0.20	0.44**
Δ HBI								0.84***	0.77***	0.11	0.19
Δ PRO2									0.89***	0.16	0.16
Δ PRO3										0.22	0.44**
Δ Fecal calprotectin											0.38*

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Δ indicates change between baseline and week 10.

BSFS, Bristol stool form scale; CDAI, Crohn's disease activity index; CRP, C-reactive protein; HBI, Harvey Bradshaw index; PRO, patient-reported outcome; VAS, visual analog scale.

the first day of the diary was 24mm (SD: 27), and mean average VAS score during week of the diary was 20mm (SD: 20).

Relationship with disease activity

The correlation of VAS scores at baseline with the clinical disease activity index scores, PRO2, PRO3, fecal calprotectin and CRP is shown in Table 2. At baseline, the VAS score correlated weakly to very weakly with clinical disease activity (r_s : 0.18–0.35) and PRO2 (r_s : 0.19–0.33) and weakly to moderately with biochemical disease activity outcomes (r_s : 0.24–0.43) and PRO3 (r_s : 0.31–0.45).

Criterion validity and responsiveness to change

The correlations between changes in VAS scores and changes in clinical disease activity index scores, fecal calprotectin and CRP after treatment are provided given in Table 3 and Fig. 3a–f. Changes in VAS scores correlated moderately to strongly with changes in clinical disease activity (r_s : 0.55–0.71), PRO2 (r_s : 0.59–0.71) and PRO3 (r_s : 0.57–0.71), whereas only a weak correlation was observed with changes in fecal calprotectin (r_s : 0.21–0.23) and CRP (r_s : 0.21–0.35).

Guyatt’s responsiveness statistics of mean VAS score for response based on CDAI and HBI, fecal calprotectin and CRP were 2.17, 1.38, 1.30 and 1.11, respectively, indicating high responsiveness. Guyatt’s statistics of VAS score at

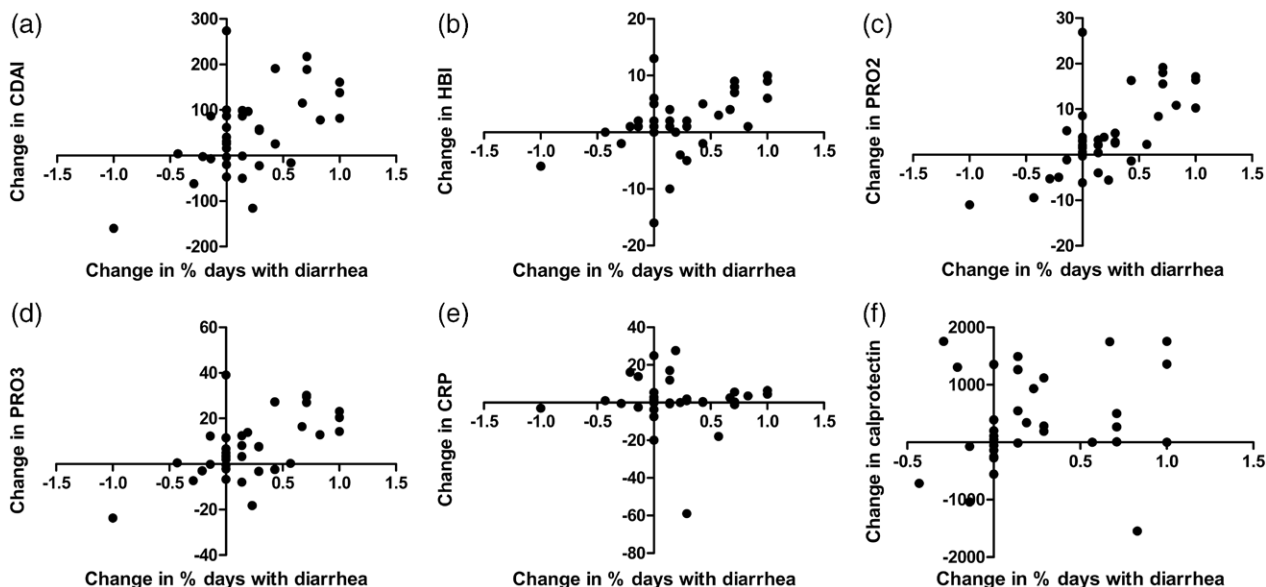


Fig. 2. Correlation between changes in the proportion of days with diarrhea based on BSFS and changes in CDAI (a), HBI (b), PRO2 (c), PRO3 (d), fecal calprotectin (e) and CRP (f). Spearman’s rank correlation coefficients are provided (r_s). BSFS, Bristol stool form scale; CDAI, Crohn’s disease activity index; CRP, C-reactive protein; HBI, Harvey Bradshaw index; PRO, patient-reported outcome.

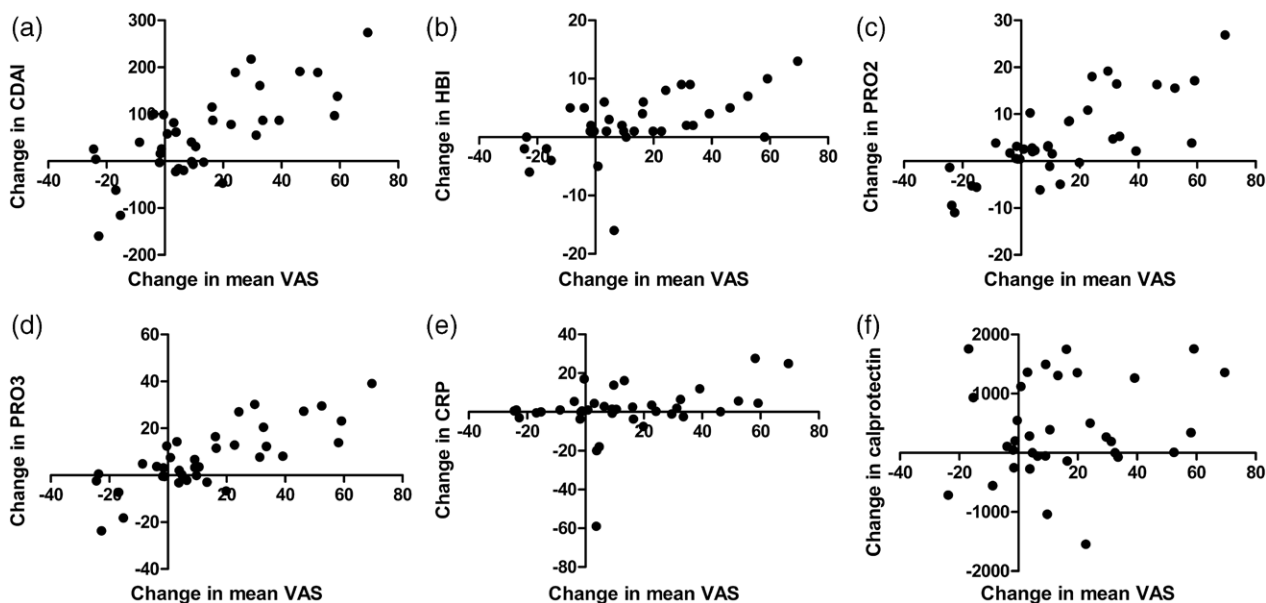


Fig. 3. Correlation between changes in mean VAS score and changes in CDAI (a), HBI (b), PRO2 (c), PRO3 (d), fecal calprotectin (e) and CRP (f). Spearman’s rank correlation coefficients are provided (r_s). BSFS, Bristol stool form scale; CDAI, Crohn’s disease activity index; CRP, C-reactive protein; HBI, Harvey Bradshaw index; PRO, patient-reported outcome; VAS, visual analog scale.

D1 were 1.59, 1.12 (high responsiveness), 0.76 and 0.62 (moderate responsiveness), respectively.

Discussion

In this prospective observational study, we evaluated the criterion validity and responsiveness of the BSFS and VAS as outcome measures for Crohn's disease in patients starting with corticosteroids or TNF inhibitors for active disease. BSFS showed a weak-to-strong correlation with various clinical disease activity indices at baseline and changes in BSFS correlated weakly to moderately with changes in clinical disease activity after treatment. VAS score (for abdominal pain) correlated weakly with clinical disease activity at baseline, while a moderate-to-strong correlation of changes in VAS score was observed with changes in clinical disease activity after treatment. Biochemical disease activity correlated weakly to very weakly with BSFS and VAS scores, both at baseline and when evaluated as change after treatment compared to baseline. Responsiveness of both BSFS and VAS was large with respect to clinical response, and moderate with respect to biochemical response.

Overall, both BSFS and VAS were more strongly related to clinical disease activity indices compared to biochemical markers. This was to be expected, considering that in both clinical activity indices abdominal pain and liquid or soft stools are important components of the total score [1,15]. For clinical practice and trials, it may therefore be preferable to score stool frequency and abdominal pain separately, with different responsiveness criteria. In reality, Crohn's disease patients with colonic inflammation have predominantly diarrhea, whereas patients with small bowel disease tend to suffer more from abdominal pain and less diarrhea. It would be interesting to separate patients with colonic and small bowel disease. Unfortunately, subgroups were too small for meaningful analyses.

Generally, changes in VAS score were more closely related to changes in both clinical and biochemical disease activity, compared to BSFS outcomes. This may be related to the fact that symptoms of diarrhea often persist in patients with Crohn's disease. Indeed, in our cohort, the vast majority of patients (74%) still suffered from diarrhea at least weekly. Considering that biochemical response was observed in roughly half of our patients, part of these symptoms of diarrhea may be explained by coexisting diarrhea-predominant irritable bowel syndrome and by bile acid diarrhea, both of which are known to occur frequently in patients with Crohn's disease [16,17].

For the BSFS and VAS, there appeared to be little advantage to determine the highest score over a seven-day period compared to using the score on the first day of the week. This may increase the feasibility and usability in both trials and clinical practice where the BSFS and VAS scores are of interest. However, the proportion of days with diarrhea derived from the diary correlated most strongly with disease activity and was found to be most responsive.

Strength of this study is the prospective collection of data in a real-life cohort of patients with Crohn's disease who require TNF inhibitors or corticosteroids for active disease. There are, however, some limitations. First, endoscopy was neither performed at inclusion nor

during follow-up. Hence, biological response was determined using the surrogate markers CRP and fecal calprotectin rather than endoscopy. Furthermore, patients were asked daily to score which types of stools they have had on the BSFS, but not the relative frequency of each individual BSFS stool type. Consequently, it was not possible to distinguish between patients who had soft or liquid stools only once during the day and those with more severe symptoms of diarrhea. Another potential limitation is that compliance with paper diaries may be low and backfilling of diaries based on recall may influence outcomes [18].

Although this study provides evidence that BSFS and VAS for abdominal pain may be used as outcome measures in patients with Crohn's disease, our results are insufficient to recommend the routine use of these instruments in clinical trials or in clinical practice. Prior to routine implementation, our findings should first be validated in an independent cohort. Based on our study, ideal BSFS and VAS candidate outcomes are (1) the proportion of days with diarrhea (i.e., BSFS ≥ 6) during 1 week, (2) one-day highest BSFS score, and (3) one-day VAS abdominal pain score.

Conclusion

The BSFS and a VAS for abdominal pain appear to be responsive instruments with moderate to strong construct validity to monitor patients with Crohn's disease. Future studies are required to validate these findings in an independent cohort.

Acknowledgements

We thank Larry Stitt for his valuable comments on the manuscript.

Conflicts of interest

G.D.H. has served as advisor for Abbvie, Ablynx, Actogenix, Amakem, Amgen, AM Pharma, AstraZeneca, Avaxia, Bristol Meiers Squibb, Boehringer Ingelheim, Celgene, Celltrion, Cosmo, Covidien, Elan, Ferring, DrFALK Pharma, Centocor/Jansen Biologics, Engene, Ferring, Galapagos, Gilead, Glaxo Smith Kline, Hospira, Medimetrics, Millenium/Takeda, Mitsubishi Pharma, Merck Sharp Dome, Mundipharma, Novonordisk, Otsuka, Pfizer, Protein Design Laboratories, Prometheus laboratories/Nestle, Receptos, Robarts Clinical Trials, Salix, Sandoz, Setpoint, Shire, Teva, Tigenix, Tillotts, Topivert, UCB, Versant and Vifor and received speaker fees from Abbvie, Ferring, Jansen Biologics, Merck Sharp Dome, Mundipharma, Norgine, Shire, Takeda, Tillotts, UCB and Vifor. C.Y.P. received grant support from Takeda, speaker's fee from Takeda, Abbvie, and Dr Falk Pharma, and consultancy fee from Takeda. M.L. received speaker fees from Abbvie, Celgene, Covidien, Dr. Falk, Ferring Pharmaceuticals, Gilead, GlaxoSmithKline, Janssen-Cilag, Merck Sharp & Dohme, Pfizer, Protagonist therapeutics, Receptos, Takeda, Tillotts and Tramedico, and has received grant support from AbbVie, Merck Sharp & Dohme, Achmea healthcare and ZonMW. G.B. is an employee of GlaxoSmithKline. There are no conflicts of interest for the remaining authors.

References

- 1 Best WR, Bechtel JM, Singleton JW, Kern F Jr. Development of a Crohn's disease activity index. National cooperative crohn's disease study. *Gastroenterology* 1976; 70:439–444.
- 2 Cellier C, Sahmoud T, Froguel E, Adenis A, Belaiche J, Bretagne JF, et al. Correlations between clinical activity, endoscopic severity, and biological parameters in colonic or ileocolonic Crohn's disease. A prospective multicentre study of 121 cases. The groupe d'études thérapeutiques des affections inflammatoires digestives. *Gut* 1994; 35:231–235.
- 3 Shah SC, Colombel JF, Sands BE, Narula N. Systematic review with meta-analysis: mucosal healing is associated with improved long-term outcomes in Crohn's disease. *Aliment Pharmacol Ther* 2016; 43:317–333.
- 4 Rutgeerts P, Vermeire S, Van Assche G. Mucosal healing in inflammatory bowel disease: impossible ideal or therapeutic target? *Gut* 2007; 56:453–455.
- 5 Levesque BG, Sandborn WJ, Ruel J, Feagan BG, Sands BE, Colombel JF. Converging goals of treatment of inflammatory bowel disease from clinical trials and practice. *Gastroenterology* 2015; 148:37–51.e1.
- 6 FDA. *Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*. 2009;(December).
- 7 Jairath V, Khanna R, Zou GY, Stitt L, Mosli M, Vandervoort MK, et al. Development of interim patient-reported outcome measures for the assessment of ulcerative colitis disease activity in clinical trials. *Aliment Pharmacol Ther* 2015; 42:1200–1210.
- 8 Khanna R, Zou G, D'Haens G, Feagan BG, Sandborn WJ, Vandervoort MK, et al. A retrospective analysis: the development of patient reported outcome measures for the assessment of Crohn's disease activity. *Aliment Pharmacol Ther* 2015; 41:77–86.
- 9 Vermeire S, Schreiber S, Sandborn WJ, Dubois C, Rutgeerts P. Correlation between the Crohn's disease activity and harvey-bradshaw indices in assessing Crohn's disease severity. *Clin Gastroenterol Hepatol* 2010; 8:357–363.
- 10 Su C, Lichtenstein GR, Krok K, Brensinger CM, Lewis JD. A meta-analysis of the placebo rates of remission and response in clinical trials of active Crohn's disease. *Gastroenterology* 2004; 126:1257–1269.
- 11 Sands BE, Chen J, Feagan BG, Penney M, Rees WA, Danese S, et al. Efficacy and safety of MEDI2070, an antibody against interleukin 23, in patients with moderate to severe Crohn's disease: a phase 2a study. *Gastroenterology* 2017; 153:77–86.e6.
- 12 Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, Spiller RC. Functional bowel disorders. *Gastroenterology* 2006; 130:1480–1491.
- 13 Swinscow TDV. *Statistics at Square One*. London: BMJ Publishing Group; 1997. <http://edmgr.ovid.com/ejgh/accounts/ifaauth.htm>.
- 14 Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000; 53:459–468.
- 15 Harvey RF, Bradshaw JM. A simple index of Crohn's-disease activity. *Lancet* 1980; 1:514.
- 16 Halpin SJ, Ford AC. Prevalence of symptoms meeting criteria for irritable bowel syndrome in inflammatory bowel disease: systematic review and meta-analysis. *Am J Gastroenterol* 2012; 107:1474–1482.
- 17 Hoekman DR, Zeevenhooven J, D'Haens GR, Benninga MA. The prevalence of irritable bowel syndrome-type symptoms in inflammatory bowel disease patients in remission. *Eur J Gastroenterol Hepatol* 2017; 29:1086–1090.
- 18 Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. *Control Clin Trials* 2003; 24:182–199.