A Noninferiority Randomized Clinical Trial of the Use of the Smartphone-Based Health Applications IBDsmart and IBDoc in the Care of Inflammatory Bowel Disease Patients

Andrew McCombie, PhD,[*] Russell Walmsley, MBChB, FRACP,[†] Murray Barclay, MD, FRACP,[‡] Christine Ho, PGDipHealSc,[§] Tobias Langlotz, PhD,[¶] Holger Regenbrecht, PhD,[‖] Andrew Gray,[**] Nideen Visesio, MN,[††] Stephen Inns, FRACP, MD(Research), MBChB,[‡‡] and Michael Schultz, MD, PhD[§§]

Background: Providing timely follow-up care for patients with inflammatory bowel disease in remission is important but often difficult because of resource limitations. Using smartphones to communicate symptoms and biomarkers is a potential alternative. We aimed to compare outpatient management using 2 smartphone apps (IBDsmart for symptoms and IBDoc for fecal calprotectin monitoring) vs standard face-to-face care. We hypothesized noninferiority of quality of life and symptoms at 12 months plus a reduction in face-to-face appointments in the smartphone app group.

Methods: Inflammatory bowel disease outpatients (previously seen more often than annually) were randomized to smartphone app or standard face-to-face care over 12 months. Quality of life and symptoms were measured quarterly for 12 months. Acceptability was measured for gastroenterologists and patients at 12 months.

Results: One hundred people (73 Crohn's disease, 49 male, average age 35 years) consented and completed baseline questionnaires (50 in each group). Intention-to-treat and per-protocol analyses revealed noninferiority of quality of life and symptom scores at 12 months. Outpatient appointment numbers were reduced in smartphone app care ($P < 0.001$). There was no difference in number of surgical outpatient appointments or number of disease-related hospitalizations between groups. Adherence to IBDsmart (50% perfect adherence) was slightly better than adherence to IBDoc (30% perfect adherence). Good acceptability was reported among most gastroenterologists and patients.

Conclusions: Remote symptom and fecal calprotectin monitoring is effective and acceptable. It also reduces the need for face-to-face outpatient appointments. Patients with mild-to-moderate disease who are not new diagnoses are ideal for this system.

Clinical Trial Registration Number: ACTRN12615000342516.

Key Words: mHealth, eHealth, remote symptom monitoring

Received for publications June 14, 2019; Editorial Decision September 12, 2019.

From the *University of Otago, Dunedin, New Zealand; †University of Auckland and Department of Gastroenterology, Waitemata District Health Board, Auckland, New Zealand; ‡University of Otago and Department of Gastroenterology, Canterbury District Health Board, Christchurch, New Zealand; §Department of Gastroenterology, Southern District Health Board, Dunedin, New Zealand; ¶Department of Information Science, University of Otago, Dunedin, New Zealand; ‖Department of Information Science, University of Otago, Dunedin, New Zealand; *Centre for Biostatistics, University of Otago, Dunedin, New Zealand; ††Department of Gastroenterology, Waitemata District Health Board, Auckland, New Zealand; ‡‡Department of Medicine, University of Otago, Wellington, New Zealand and Department of Gastroenterology, Hutt Valley District Health Board, Lower Hutt, New Zealand; §§Department of Medicine, University of Otago, and Department of Gastroenterology, Southern District Health Board, Dunedin, New Zealand

Supported by: This work was supported by The Healthcare Otago Charitable Trust (no grant number) and The New Zealand Society of Gastroenterology Janssen Research Fellowship (no grant number) in 2015 and the guthealthnetwork, a research theme located at the Department of Medicine, University of Otago (www.guthealthnetwork.com).

Conflicts of interest: None to declare.

Address correspondence to: Andrew McCombie, BSc, BA(Hons), PhD, Department of Medicine, University of Otago, PO Box 4345, Christchurch 8140, New Zealand (mccombieandrew@hotmail.com).

doi: 10.1093/ibd/izz252
Published online 23 October 2019

INTRODUCTION

Inflammatory bowel disease (IBD), which includes Crohn's disease (CD) and ulcerative colitis (UC), is characterized by chronic relapsing–remitting inflammation of the gastrointestinal tract, in particular the small and large intestines. Symptoms include diarrhea, rectal bleeding, abdominal pain, and extra-intestinal manifestations (eg, arthritis, liver symptoms, skin and eye irritation).[1] Management includes medication that often requires close monitoring, along with surgery and dietary approaches. Health-related quality of life (HRQOL) is normal in those in remission,[2,3] so obtaining and maintaining remission are the treatment goal.

Given the chronicity of IBD, many patients require lifelong care under a gastroenterologist to manage their disease. However, with IBD prevalence rising[4,5] and physician shortages regularly reported,[6] more cost- and time-efficient management approaches to IBD management are called for. The present model of care often leads to asymptomatic patients being seen in outpatient clinics at the expense of symptomatic patients. Inflammatory bowel disease management efficiencies may be improved through self-management approaches. For example,

a large multicenter cluster randomized controlled trial (RCT) performed in the United Kingdom reported that a self-management approach reduced hospital visits without increasing the number of primary care visits and without a difference in HRQOL or anxiety.[7] Moreover, a recent Dutch study reported the telemedicine group to have fewer outpatient appointments and hospital admissions than the standard care group.[8]

The advent of smartphone apps has provided a new and potentially revolutionary means for IBD patients to manage their illness and communicate with their health care provider. However, smartphone apps for IBD to date have generally lacked physician input, have not used clinically validated indices, and have not been clinically tested to see whether they improve HRQOL, reduce the need for face-to-face (F2F) appointments, or enhance cost-effectiveness apart from the subjective app ratings and reviews in the stores.[9, 10] As stated by Spiegel, "Although enthusiasm for mHealth is boiling over, the level of evidence does not match the level of excitement."[11]

This study was performed to help reverse the trend of smartphone apps being developed for IBD patients but not tested. This study simultaneously tested 2 smartphone apps for IBD monitoring and management, namely IBDsmart and IBDoc. IBDsmart has the core function of recording symptom disease activity index scores via the well-validated Harvey Bradshaw Index (HBI)[12] and the Simple Clinical Colitis Activity Index (SCCAI).[13] These symptom scores can be sent to health care providers from the app at any time. The app has been piloted on 35 IBD patients,[14] and overall 74% of the responses were clearly positive with respect to the general usability of the system. Since this pilot, the app has been further developed for use in the present study. The advantages of IBDsmart are its simplicity and lack of clutter (relative to a system like myIBDcoach), it is not associated with pharmaceutical industry funding, and it uses validated scoring systems.

IBDoc has the core function of producing fecal calprotectin (FC) scores from stool samples provided by the IBD patient in their own home.[15, 16] Fecal calprotectin is a biomarker that can be used to predict IBD relapse,[17] so it has great potential as a self-monitoring tool. Encouragingly, testing in Europe has not only demonstrated high usability of IBDoc[16] but also very high validity; the performance of IBDoc is similar to that of laboratory-based methods.[15] For the purpose of this study, the information provided by IBDoc augmented the clinical symptom–based data produced by IBDsmart.

The primary aim of this study was to test whether IBDsmart and IBDoc are noninferior to standard care in terms of HRQOL and IBD symptoms (measured via HBI or SCCAI) at 52 weeks and whether those in the smartphone app group had fewer F2F appointments compared with those in standard care over the 52 weeks. Secondarily, the usability and acceptability of IBDsmart and IBDoc were measured for physicians and patients at other time points (3, 6, and 9 months). It was hypothesized that IBDsmart and IBDoc would reduce outpatient appointments without leading to a deterioration in HRQOL or symptoms and that the apps would be usable and acceptable to gastroenterologists and patients.

METHODS

Trial Design

This study was a 52-week prospective, multicenter, noninferiority RCT comparing IBDsmart/IBDoc-assisted virtual clinic appointments (intervention) with routine F2F clinic appointments (control). Upon randomization, all participants completed baseline questionnaires and F2F HBIs or SCCAIs. Participants in both groups completed inflammatory bowel disease questionnaires (IBDQs) 3-monthly starting from baseline. Smartphone app patients also completed HBIs or SCCAIs 3-monthly starting from baseline, whereas standard care patients completed HBIs or SCCAIs at F2F clinics. At 12 months, all participants had F2F HBIs or SCCAIs again.

Participants

The participants were recruited from the Southern, Canterbury, Waitemata, and Hutt Valley District Health Boards across New Zealand.

Inclusion/Exclusion Criteria

Included participants were those with confirmed UC or CD, who had at least 2 outpatient appointments in the last 12 months, had <3 disease flares in the past 12 months, were willing and able to provide written consent, and were aged 16 years or older. Excluded participants were those with indeterminate colitis, with severe disease requiring close monitoring, with possible/planned surgical intervention forthcoming, with an ileostomy, colostomy, or ileal pouch–anal anastomosis, and who were pregnant.

Recruitment Methods and Study Settings

Participants were recruited from multiple gastroenterologists' outpatient appointments from August 6, 2015, until December 23, 2016. If they were interested and eligible, they were approached and given an information sheet and consent form. Those who opted out were not contacted further.

Interventions

After consent was obtained and patients were stratified by disease type and location of outpatient appointments, participants were randomized to smartphone app or standard care. Those who were allocated to the smartphone app group were given instructions on how to use the apps and received a username and password for the study. Once the induction was

completed, all participants in both groups were given their baseline questionnaires to take home and send back via mail; they were given a choice to complete subsequent questionnaires via an online survey or pen-and-paper. Participants were not blinded to which group they were in.

Smartphone app group

Those in the smartphone app group had access to the apps, namely IBDsmart and IBDoc. IBDsmart is a symptom monitoring app that uses the HBI[12] and SCCAI[13] to monitor the disease activity of IBD patients. It should be noted the primary measures requested by participants in the present study were HBI for CD patients and SCCAI for UC patients. Upon completion, the symptom scores are sent to the treating gastroenterologists. There is also the possibility to report a flare through IBDsmart. A comment box is provided at the end of the clinical index to allow patients to report anything outside the realm of what the HBI or SCCAI asks about (eg, pregnancy, side effects of medications, etc.).

IBDoc measures FC levels from IBD patients' stool samples in their own home.[15, 16] It includes a small piece of equipment for reading FC levels that produces an output that can be read by the camera on the smartphone. IBDoc then produces a number ranging from <30.0 µg/g to >1000.0 µg/g. Again, the results from this can be sent directly to the patient's health care team.

It should be noted that patients were strongly encouraged to complete their symptom scores and FC levels at baseline and at 3, 6, 9, and 12 months, irrespective of how well they felt, but they were also encouraged to complete these if feeling unwell between these time points.

Standard care group

Those in the standard care group received their usual IBD care from their physician. Both groups completed questionnaires at baseline and at 3, 6, 9, and 12 months after baseline.

Outcomes

Primary

Noninferiority. The primary outcome was the noninferiority of IBDsmart and IBDoc to standard care. In order for noninferiority to be determined, the patients in the smartphone app group would have to have noninferior HRQOL and symptoms compared with the standard care group at 12 months. Noninferiority was measured both per-protocol and intention-to-treat (ITT). This noninferiority was expected, in addition to a reduction in F2F outpatient appointments.

HRQOL was measured using the IBDQ.[18] The IBDQ contains 32 items divided into 4 health subdimensions: bowel symptoms (eg, loose stools, abdominal pain; 10 items), systemic symptoms (eg, fatigue, sleeping problems; 5 items), social functioning (eg, limited social activity, school or work attendance; 5 items), and emotional function (eg, irritability, anger, depression; 12 items). Responses are scored on a 7-point scale where 7 corresponds to the best function and 1 to the worst. A study of UC patients found that the IBDQ has an SD of 48.00 and a clinically significant change score is 20.[19] Participants in both groups answered the IBDQ at baseline and at 3, 6, 9, and 12 months.

Symptoms of UC patients were measured using the SCCAI,[13] which consists of 5 clinical questions. A higher score corresponds to worse symptoms. Scores of ≤2 indicate remission and ≥3 indicate relapse.[19]

Symptoms of CD patients were measured using the HBI.[12] The HBI was devised as a simpler and more concise version of the Crohn's Disease Activity Index (CDAI).[12] On the HBI, a cutoff of ≤4 is considered remission, whereas >4 is considered a relapse.[20] It correlates well with the more complex CDAI.[12, 20]

Patients in the standard care group completed the HBI/SCCAI via clinical interview at all appointments between baseline and 12 months, whereas patients in the smartphone app group completed the HBI/SCCAI at least quarterly in a self-report manner in between baseline and 12 months (ie, 3, 6, and 9 months). Smartphone app patients had the HBI/SCCAI completed via clinical interview and via IBDsmart at baseline and 12 months.

Secondary

Quality of life and symptoms at 3, 6, and 9 months. Secondarily, HRQOL and symptoms were tested for noninferiority at 3, 6, and 9 months using the IBDQ,[18] SCCAI (for UC),[13] and HBI (for CD).[12]

Patient-reported usability/acceptability. At the end of 12 months, patients in the smartphone app group completed 2 system usability scales (SUS),[21] 1 for IBDsmart and 1 for IBDoc. This is a 10-item questionnaire answered on a 5-point scale. Scores can range from 0 to 100, and higher scores indicate better usability of the system. Patients also completed usability questionnaires developed for the present study about IBDsmart and IBDoc and 1 questionnaire directly comparing the 2 apps (Supplementary Data). The questionnaires asked about the instructions provided for the apps, what issues with the apps they experienced during the study, and whether they would keep using the apps in the future and recommend them to other people with IBD. Participants were also asked questions to directly compare the apps in terms of 10 different attributes and overall usability.

Doctor-repored usability/acceptability. For each patient at 12 months, their gastroenterologist completed an acceptability questionnaire (Supplementary Data). The questions were worded differently according to which group the patient was in. The doctor was asked about comfort and whether enough

information was provided about the patient in their F2F or smartphone app–based consultations. They were also asked if there was anything they were unable to communicate via the consultation method.

Adherence. Adherence was measured by recording how many patients completed the apps at each time point and how much contact was required to get the patients to complete the apps at each time point.

Fecal calprotectin. FC was requested via IBDoc from smartphone app patients at baseline and at 3, 6, 9, and 12 months. FC scores were analyzed at these time points, although no FC scores were provided from the F2F group.

Randomization

Randomization occurred by a computer program randomly allocating participants to 1 of the 2 groups. Randomization was stratified by disease type (CD vs UC) and location of outpatient appointments (Waitemata, Hutt Valley, Canterbury, and Southern District Health Boards). The allocations were put in sequenced envelopes, which were to be opened by the recruiting nurse, gastroenterologist, or researcher.

Participant Contact via Email, Text Message, and Phone Call

Participants were contacted if they had not completed their IBDQ or 1 of the 2 smartphone apps, which was at least once every 3 months over 1 year. Records were kept of the number of contacts from the researcher required for completion of each IBDQ, IBDsmart recording, or IBDoc reading. Participants were systematically contacted by the least invasive method first, namely email. If they still needed reminding, they were then contacted via text message. Finally, participants were contacted by phone if they still had an aspect of the study to complete.

Statistical Methods

Power calculations determined that a sample size of 31 patients with CD per group (n = 62) at follow-up would provide 80% power to detect noninferiority ($P < 0.05$) using the HBI, assuming an SD of 4.7 and an equivalence limit of 3. In addition, a sample size of 17 patients with UC per group (n = 34) at follow-up would provide 80% power to detect noninferiority ($P < 0.05$) using the SCCAI, assuming an SD of 3.5 and an equivalence limit of 3. Finally, a sample size of 45 patients with either CD or UC per group (n = 90) at follow-up would provide 80% power to detect noninferiority ($P < 0.05$) using the IBDQ, assuming an SD of 38 and an equivalence limit of 20. Thus, the study was adequately powered, with a total of 96 participants (62 with CD and 34 with UC) at follow-up. Allowing for 5% loss to follow-up, the aim was to recruit 102 participants (66 with CD and 36 with UC) at baseline.

The data were managed by the study administrator, and all patient information was partially de-identified. Appropriate descriptive statistics were provided for all measures of interest. For the noninferiority questions, IBDQ, SCCAI, and HBI scores (with noninferiority limits of –20, –3, and –3, respectively) were compared between groups using linear mixed models to model differences in changes between groups, adjusting for baseline scores along with stratification variables at 3, 6, 9, and 12 months, with the last of these times being the primary end point. Noninferiority was determined using 90% confidence intervals at each time point (equivalent to a 1-sided test at the 0.05 level). Model diagnostics included visual checks of normality and homoscedasticity for residuals. Both ITT and per-protocol analyses were performed, with noninferiority concluded when both models showed evidence of noninferiority. Poisson regression was used to compare health service use (counts of F2F gastroenterologist and surgical appointments and number of IBD-related hospitalizations and associated nights in the hospital) during the study period, with a likelihood ratio test for overdispersion to identify when negative binomial regression was more appropriate. These models adjusted for stratification variables only, and a 2-sided P value <0.05 was considered statistically significant. Where the number of 0 counts prevented this approach, exact Wilcoxon rank-sum tests were used instead. The analyses described in this paragraph were performed using Stata 15.1, and the statistician was blinded to which groups the participants were in until all planned statistical analyses were completed. No formal plan was made for missing data, nor were any adjustments made for multiple comparisons.

Analyses were also performed on attrition (percentages), adherence (percentages), usability (means and SDs), physician-reported acceptability (percentages), and fecal calprotectin (boxplots). Chi-square analyses were also performed to compare the number of HBI/SCCAI-defined flares in smartphone apps vs standard care. These specific analyses were performed using Statistical Packages for Social Sciences, version 25.[22]

Ethical Considerations

The study was conducted with the consent of the New Zealand Health and Disability Ethics Committee (15/NTA/44). Each participant provided informed consent, and only de-identified data are presented. The study was registered with the Australian New Zealand Clinical Trials Registry (ACTRN12615000342516).

RESULTS

Participant Flow

Participants were recruited from the Southern, Canterbury, Waitemata, and Hutt Valley District Health Boards. The participant flow is shown in Figure 1. Of the 107 randomized participants, 7 patients were excluded after randomization, as they did not complete the baseline questionnaire. Forty-three (86%) of

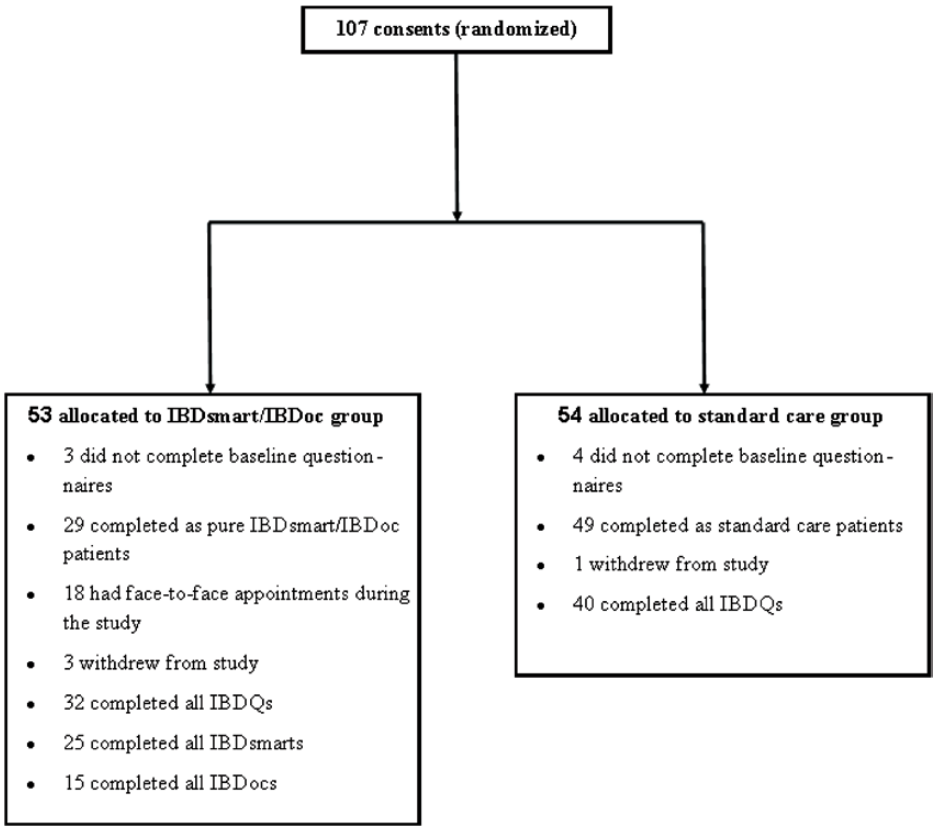

FIGURE 1. Participant flow.

the smartphone app participants used both apps at least once. Participants were recruited from August 6, 2015, to December 23, 2016. The last participant had their 52-week follow-up performed on December 23, 2017.

Baseline Data

Baseline data from both groups are shown in Table 1. There were no baseline differences between the groups.

Attrition

At baseline, 50 (94%) smartphone app and 50 (93%) standard care participants completed the questionnaires. By the end of the 12 months of follow-up, there were 47 (89%) smartphone app patients still in the study and 49 (91%) in the standard care group.

Outcomes and Estimation

Quality of life

For all IBD patients, HRQOL as measured by the IBDQ was noninferior at all time points in both ITT and per-protocol analyses (Table 2; Supplementary Figs 1 and 2). When patients were separated into CD and UC, only UC patients at 3 months in ITT analyses were shown to be inferior (Table 2).

Symptom scores

SCCAI. For UC patients, the SCCAI was noninferior at all time points in ITT analyses and per-protocol analyses (Table 2; Supplementary Figs 3 and 4). Seventy percent of smartphone app patients had reported an SCCAI flare score at months 3–12, compared with 41.7% of standard care patients ($\chi^2 = 1.77$; $P = 0.18$).

HBI. For CD patients, the HBI was noninferior at all time points except for 9 months (Table 2; Supplementary Fig. 5) in ITT analyses. Per-protocol analyses showed noninferiority at all time points except for 6 and 9 months (Table 2; Supplementary Fig. 6). Forty-seven point two percent of smartphone app patients reported an HBI flare score at months 3–12, compared with 25.7% of standard care patients ($\chi^2 = 3.54$; $P = 0.06$).

Health care use. ITT and per-protocol analyses revealed that smartphone app patients had less gastroenterology F2F appointments than standard care patients and there were no differences in other health care usage outcomes (Table 3).

Patient-reported usability of apps

IBDsmart. Of the 47 smartphone app patients who did not dropout during the study, 31 participants answered the SUS (66%). The mean SUS (SD) was 81.4 (14.1), indicating high usability. Figure 2 summarizes the answers to the other usability

TABLE 1. Baseline Demographics, Disease-Related Variables, and Outpatient Care Answers (all People who Consented and Completed Baseline Questionnaires)

	Smartphone App Care	Standard Care
	n = 50	n = 50
Variable	No. (%)/Mean (SD)	No. (%)/Mean (SD)
Crohn's disease	37 (74%)	36 (72%)
Male	26 (52%)	23 (46%)
Age, y	35.2 (12.4)	34.3 (12.9)
Years since diagnosis	7.7 (7.0)	9.5 (8.5)
District Health Board		
Southern	13 (26%)	12 (24%)
Canterbury	18 (36%)	19 (38%)
Hutt Valley	7 (14%)	5 (10%)
Waitemata	12 (24%)	14 (28%)
Inflammatory Bowel Disease Questionnaire	177.1 (26.3)	173.2 (31.8)
Simple Clinical Colitis Activity Index (Interview)	1.6 (2.5)	1.1 (1.5)
Harvey-Bradshaw Index (Interview)	2.7 (3.00)	2.7 (3.0)
Medications		
5-ASA	20 (40%)	20 (40%)
Biological	15 (30%)	18 (36%)
Thiopurine or methotrexate	37 (74%)	27 (54%)
None	2 (4%)	3 (6%)
Borrowed phone for the study	17 (34%)	Not applicable
No. outpatient appointments in the last 12 mo	2.9 (1.3)	3.0 (1.5)
No. IBD nurse face-to-face contacts in the last 12 mo	1.1 (1.2)	1. 6 (1.6)
How do you usually make contact with the IBD nurse?		
Email	9 (18%)	9 (18%)
Text	9 (18%)	13 (26%)
Mobile call	19 (38%)	21 (42%)
Landline call	16 (32%)	9 (18%)
More than 1	10 (20%)	12 (24%)
None or not applicable	10 (20%)	9 (18%)
Other	0 (0%)	3 (6%)
How good are you at using a smartphone?		
OK	11 (22%)	Not applicable
Good	20 (40%)	Not applicable
Excellent	19 (38%)	Not applicable
How do you usually get to outpatient appointments?		
Car	41 (82%)	45 (90%)
Walk	4 (8%)	2 (4%)
More than 1 at the same time (eg, car and bus)	2 (4%)	1 (2%)
Bus	2 (4%)	1 (2%)
Motorcycle	1 (2%)	0 (0%)
Cycle	0 (0%)	1 (2%)
How many minutes does it take to get to outpatient appointments each time	25.9 (19.0)	31.8 (47.1)

questions. More than 80% of respondents answered affirmatively to the questions about whether they would keep using IBDsmart in the future and whether they would recommend IBDsmart to other people.

IBDoc. Thirty-five participants answered the SUS about IBDoc (74.5% response rate). The mean SUS (SD) was 71.6 (16.8), indicating relatively good usability. Figure 2 summarizes the answers to the other usability questions. Just over half of

TABLE 2. Intention-to-Treat and Per-Protocol Analyses for IBDQ, SCCAI, and HBI

	Intention-to-Treat			Per-Protocol		
Time	Raw Mean (SD) for IBDsmart/ IBDoc	Raw Mean (SD) for F2F Control	Adjusted Difference in Changes for IBDsmart/IBDoc (90% CI)	Raw Mean (SD) for IBDsmart/ IBDoc	Raw Mean (SD) for Control	Adjusted Difference in Changes for IBDsmart/IBDoc (90% CI)
IBDQ combined UC and CD						
0	177.1 (26.3)	173.2 (31.8)		182.5 (22.4)	174.9 (31.6)	
3	176.5 (28.4)	167.7 (33.8)	4.1 (–3.6 to 11.9)	183.0 (25.3)	169.7 (32.8)	6.8 (–2.0 to 15.5)
6	179.9 (28.1)	166.7 (35.4)	8.6 (0.8 to 16.5)	186.2 (27.5)	167.3 (35.6)	9.9 (0.9 to 18.8)
9	179.5 (28.0)	165.4 (31.9)	10.2 (2.3 to 18.1)	189.0 (24.9)	166.2 (32.3)	15.3 (6.2 to 24.3)
12	181.1 (22.0)	170.6 (30.8)	4.8 (–3.3 to 12.9)	186.6 (21.9)	171.8 (31.0)	6.7 (–2.5 to 16.0)
IBDQ UC only						
0	188.1 (24.5)	185.2 (19.1)		184.9 (29.0)	188.7 (14.6)	
3	184.6 (21.7)	186.6 (21.0)	**–6.9 (–22.7 to 8.9)**	184.9 (22.6)	188.8 (20.1)	–0.2 (–16.2 to 15.9)
6	188.0 (28.6)	175.5 (31.8)	6.9 (–9.4 to 23.2)	184.8 (36.4)	178.3 (31.5)	10.7 (–6.1 to 27.6)
9	181.6 (30.4)	181.9 (27.7)	–0.1 (–16.7 to 16.5)	196.8 (15.0)	184.0 (27.9)	24.3 (6.7 to 41.8)
12	189.5 (24.5)	179.6 (24.3)	3.4 (–13.1 to 19.9)	200.5 (16.3)	183.0 (22.4)	21.0 (4.0 to 38.0)
IBDQ CD only						
0	173.3 (26.1)	168.6 (34.6)		181.7 (20.7)	169.6 (34.8)	
3	173.9 (30.0)	160.1 (35.1)	8.3 (–0.5 to 17.1)	182.4 (26.5)	162.2 (34.0)	9.2 (–1.1 to 19.6)
6	177.5 (27.9)	163.1 (36.7)	9.3 (0.4 to 18.3)	186.6 (25.5)	163.1 (36.7)	9.5 (–1.0 to 20.0)
9	178.9 (27.8)	159.0 (31.4)	13.5 (4.6 to 22.4)	187.1 (26.7)	159.7 (31.7)	11.8 (1.3 to 22.3)
12	178.0 (20.6)	167.3 (32.6)	5.4 (–3.9 to 14.7)	181.7 (21.8)	167.8 (33.0)	1.6 (–9.4 to 12.5)
SCCAI (UC only)						
0	1.4 (2.4)	1.1 (1.5)		1.7 (3.0)	1.2 (1.5)	
3	1.6 (1.7)	0.5 (0.7)	0.5 (–1.8 to 2.9)	1.5 (1.6)	0.5 (0.7)	0.3 (–2.2 to 2.8)
6	2.5 (2.2)	1.9 (2.0)	–0.3 (–1.9 to 1.4)	3.0 (2.5)	1.9 (2.0)	–0.3 (–2.2 to 1.6)
9	3.4 (2.7)	2.6 (4.8)	0.2 (–1.6 to 2.0)	2.0 (1.9)	2.6 (4.8)	–1.3 (–3.3 to 0.7)
12	1.5 (1.1)	1.7 (1.9)	–0.8 (–2.4 to 0.8)	1.2 (1.1)	1.7 (1.9)	–1.5 (–3.3 to 0.4)
HBI (CD only)						
0	2.8 (3.2)	2.7 (3.0)		1.9 (2.6)	2.6 (3.0)	
3	4.3 (3.5)	3.6 (2.3)	1.0 (–0.4 to 2.5)	3.2 (2.4)	3.6 (2.3)	1.1 (–0.3 to 2.4)
6	4.2 (3.8)	2.5 (3.1)	1.1 (–0.3 to 2.4)	3.7 (4.0)	2.5 (3.1)	**1.7 (0.4 to 3.0)**
9	3.9 (4.0)	1.8 (1.9)	**2.0 (0.5 to 3.6)**	3.1 (2.9)	1.8 (1.9)	**2.1 (0.6 to 3.5)**
12	2.4 (3.4)	2.0 (2.5)	0.3 (–1.0 to 1.5)	1.6 (2.2)	1.8 (2.4)	0.6 (–0.6 to 1.8)

Inferiority is shown in bold. For IBDQ: All 90% CI limits fall within the 20-unit noninferiority margin except for UC patients at 3 months. For SCCAI: All 90% CI limits fall within the 3-unit noninferiority margin. For HBI: All 90% CI limits fall within the 3-unit noninferiority margin, except at 9 months in intention-to-treat and 6 and 9 months in per-protocol, where IBDsmart is worse than F2F control.

respondents answered that they would like to keep using IBDoc in the future and would recommend it to other people.

IBDsmart and IBDoc comparison. Thirty-six participants completed the comparison questionnaire. Thirty-one (86.1%) agreed or strongly agreed that using IBDsmart and IBDoc together seemed logical to them. Thirty-four (94.4%) agreed or strongly agreed that they understood the importance of providing FC stool tests (via IBDoc) on top of symptom scores (via IBDsmart) to their IBD care team. The direct comparisons of the usability characteristics are shown in Supplementary Table 1. IBDsmart was considered more usable than IBDoc in terms of instructions, ease of use, time taken, fewer software problems, not crashing, getting results from the phone to the physician or IBD nurse, and overall. IBDoc was considered better in terms of improving disease self-management, accessing the score history, and replacing F2F appointments. The help function was considered equally as good in both apps.

TABLE 3. Intention-to-Treat and Per-Protocol Analyses of Health Care Usage Between Groups Differences

	Intention-to-Treat				Per-Protocol			
Outcome	Raw Mean (SD) for IBDsmart/ IBDoc	Raw Mean (SD) for Control	Ratio (Combined UC and CD) for IBDsmart/ IBDoc (95% CI)[a]	*P*	Raw Mean (SD) for IBDsmart/ IBDoc	Raw Mean (SD) for Control	Ratio (Combined UC and CD) for IBDsmart/ IBDoc (95% CI)	*P*
Gastro appointments between baseline and 12 mo	0.6 (0.9)	1.7 (0.8)	0.36 (0.24 to 0.55)	**<0.001**	0.0 (0.0)	1.7 (0.7)	—	**<0.001**
Surgical appointments between baseline and 12 mo	0.1 (0.4)	0.1 (0.4)	—	0.729	0.1 (0.4)	0.1 (0.4)	—	0.600
No. IBD-related hospitalizations during study period	0.1 (0.3)	0.1 (0.4)	—	0.473	0.0 (0.2)	0.1 (0.4)	—	0.384
No. nights in hospital due to IBD during study period	0.1 (0.4)	0.8 (3.9)	—	0.629	0.1 (0.4)	0.3 (1.1)	—	0.552

P values in bold are statistically significant at the 2-sided 0.05 level.
[a]From Poisson regression: Too few non-0 values to model using regression, *P* value from exact Wilcoxon rank-sum test.

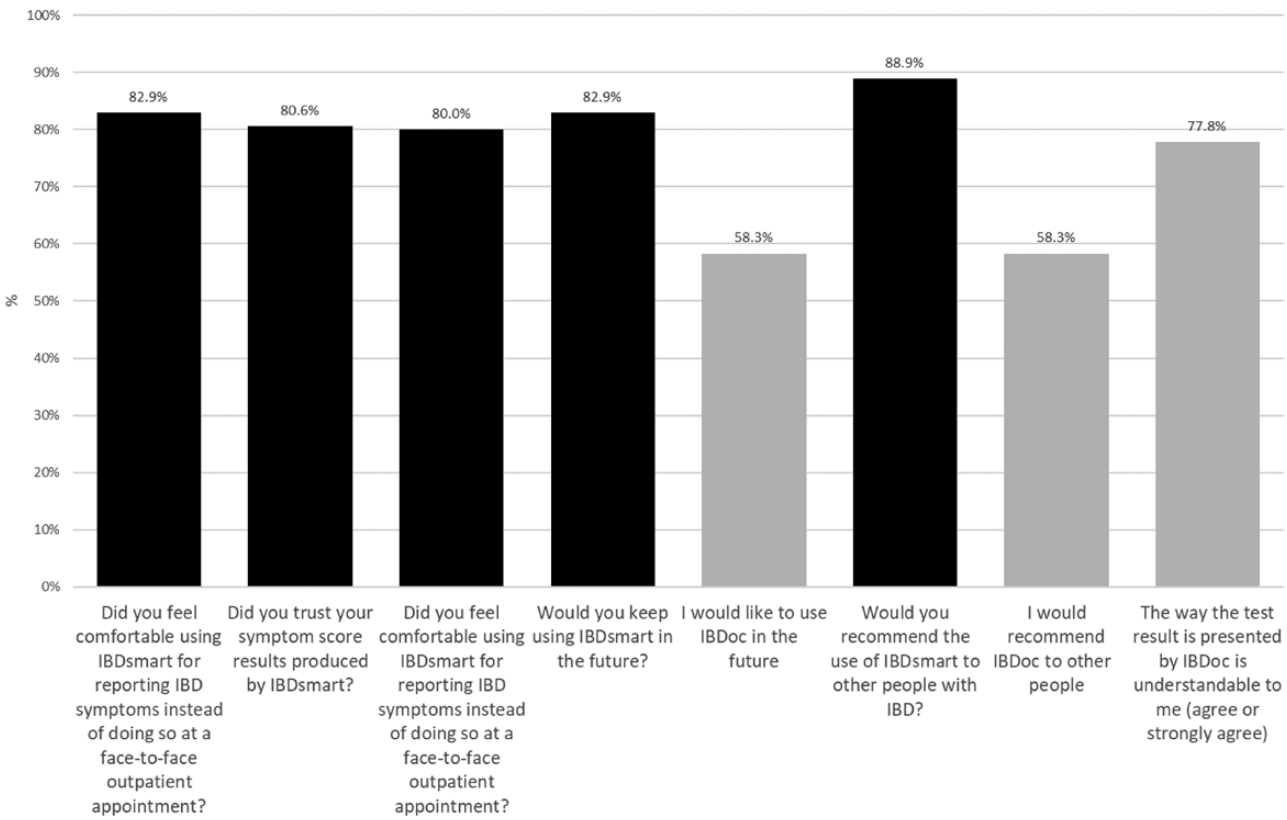

FIGURE 2. Patient-reported usability of IBDsmart and IBDoc.

Physician-reported acceptability of apps

Eleven physicians completed usability questionnaires for each participant. Supplementary Table 2 reports on the physician-reported usability for each patient they treated in the study and compares the 2 groups. Of note is that 60% of responses in the smartphone app group indicated that there was something they were not able to communicate with their patient via the smartphone apps, whereas just 12% reported the same for F2F appointments.

Adherence to apps

Adherence to IBDsmart and IBDoc is shown in Figure 3. For IBDsmart, 25 (50%) completed all readings, 9 (18%)

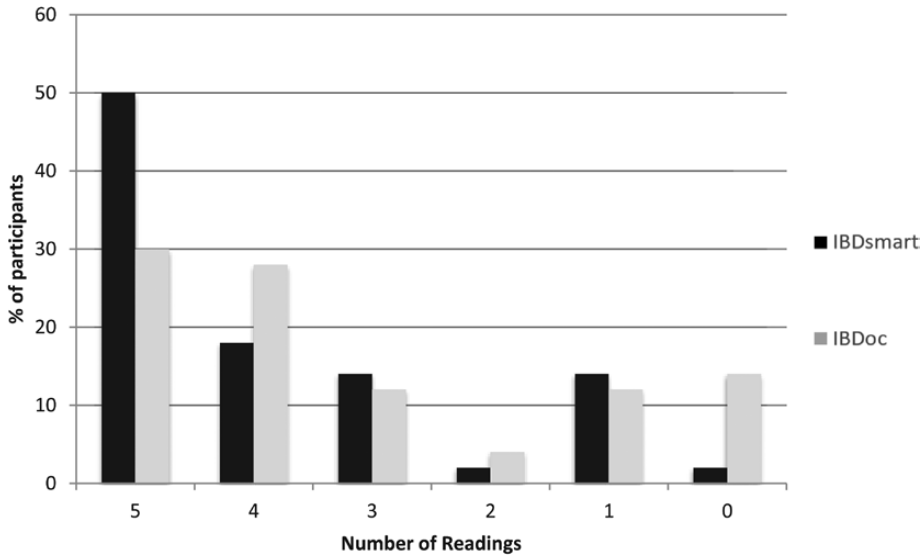

FIGURE 3. Adherence to IBDsmart and IBDoc.

completed 4, 7 (14%) completed 3, 1 (2%) completed 2, 7 (14%) completed 1, and 1 (2%) completed 0. Reasons patients reported not completing all IBDsmart indices varied (Supplementary Table 3). For IBDoc, 15 (30%) completed all readings, 14 (28%) completed 4, 6 (12%) completed 3, 2 (4%) completed 2, 6 (12%) completed 1, and 7 (14%) completed 0. Reasons patients reported not completing all IBDoc indices varied greatly (Supplementary Table 4).

Ancillary Analyses

Fecal calprotectin

Box plots of IBDoc readings (per protocol) are shown in Supplementary Figure 7. FC remained stable throughout the study for these participants, and the median was between medium and high at all time points.

DISCUSSION

This study demonstrated that IBDsmart and IBDoc reduced the number of outpatient appointments while not leading to a deterioration in symptoms or HRQOL. The primary outcome of noninferiority of HRQOL and symptoms in ITT and per-protocol analyses at 12 months was achieved. IBDsmart and IBDoc also demonstrated good usability and acceptability among patients and gastroenterologists, although IBDsmart was considered a little easier to use among patients. Although physicians did report adequate acceptability of the apps relative to standard care, a higher proportion of responses of gastroenterologists in the smartphone app group indicated that there were some things they were not able to communicate with the patient because they were seeing them via IBDsmart/IBDoc and not F2F (eg, non-IBD issues).

Implications and Limitations

Overall, in light of the need for efficient use of clinical resources in gastroenterology, the findings from this study are encouraging and are in line with other reported findings.[7,8] Outpatient appointments have been reduced by remote symptom and FC monitoring without leading to worse outcomes for the patient. This has implications for gastroenterologists with limited capacity in outpatient clinics and also has implications for patients, especially those who live far away from the hospital or have issues physically attending outpatient appointments. IBDsmart and IBDoc have allowed for resourceful triaging of outpatients. This veers away from the traditional model of estimating when the next outpatient appointment is due and more toward an "as-needed" care model. As Richardson et al. stated, "Periods of activity are unlikely to coincide with outpatient appointments when a traditional fixed appointment scheme is used. This results in unnecessary attendances for those in remission and/or a lack of access for those with disease activity."[23] IBDsmart and IBDoc can assist in prioritizing patients into "low" and "high" needs before their next appointment, thus ensuring that those with the highest needs are seen by their gastroenterologist first.

The recent Dutch study of myIBDcoach had similar findings,[8] but they used a more complicated regimen in their telemedicine arm involving not just symptom monitoring but also questions about side effects and adherence, e-learning modules, and a personal care plan. In contrast, the present study used simple symptom and FC monitoring without any psychoeducational components. Hence, it is possible that simpler telemedicine programs are noninferior to more complicated telemedicine programs, although this question needs further exploration in future research.

Many apps have been developed for IBD but have weaknesses, including a lack of clinically validated indices, heavy industry involvement, and a lack of physician input in their development.[9, 10] Conversely, the present study did use validated clinical indices, and it is important to note that HBI[24] and SCCAI[25] have been used in a self-report context before this study. In addition, physicians and patients had input into IBDsmart's development and added remote FC as an objective marker of disease monitoring, which allowed inorganic causes of clinical symptoms (eg, irritable bowel syndrome) to be ruled out when interpreting clinical index scores.

A perhaps controversial future step in telemedicine could be the development of a program for medication self-management with algorithms for determining what dose the patient should take based on their latest symptom and FC scores. A review by Jackson and de Cruz[26] summarized evidence for treating mild to moderate UC with the purpose of developing clinical algorithms that guide shared decision-making and facilitate self-management. Such algorithms could potentially be used in the future in conjunction with disease activity indices.

Despite the promising findings, there were some limitations with the use of the smartphone apps. There were some dropout and adherence issues, and it would seem some patients are not suited to remote symptom monitoring. Moreover, the number of people who declined to participate was not recorded, so it is not known what percentage of patients eligible for this study are suitable for such care. Some of the adherence problems were caused by technical glitches with the apps. Contact by the study co-ordinator with the patient via email, text, or phone call (in that order) was required in more patients as the study progressed; it seems that such a program will never be completely self-guided by the patient.

IBDoc had the added problem of being a stool test, which is less acceptable than the simple brief questionnaire that was required for IBDsmart. Nevertheless, patients seemed to understand the importance of FC monitoring, given that more than 94.4% agreed or strongly agreed that they understood the importance of providing FC stool tests (via IBDoc) on top of symptom scores (via IBDsmart) to their IBD care team, and only 19.4% agreed or strongly agreed that symptom scores (via IBDsmart) are more important than FC scores (via IBDoc). Regarding FC generally, how it is best used to monitor disease is also continuing in its development, but of note is that FC correlates well with endoscopic and histological measures of disease,[27] and the latest European Crohn's and Colitis Organisation guidelines recommend using FC to monitor IBD.[28]

Noninferiority was not found in per-protocol analyses at 6 or 9 months for HBIs or at 9 months for ITT analyses. Moreover, more flares were reported in the smartphone app group compared with standard care, although these increased flare report rates were not statistically significant. It is uncertain whether the temporary lack of noninferiority in HBI scores was real or an artefact of self-report vs clinical interview. The latter is somewhat likely given that crude analyses of self-report vs clinical interviews demonstrated that self-report scores were higher, albeit insignificantly, than clinician-completed scores. In addition, analyses of the IBDQ with CD patients only demonstrated noninferiority in CD patients. Regarding the difference in flare rates, it is possible that these are a natural consequence of smartphone app patients having ease of access to flare reporting between the 3-monthly indices, whereas the standard care patients did not have as readily available access to flare reporting between clinic visits.

There are concerns surrounding the generalizability of the findings. First, the study was performed in New Zealand at centers with specialist IBD nurses, who were heavily involved. As to whether this system can be replicated in centers without these important health care professionals remains to be seen. Second, IBDsmart and IBDoc are appropriate to be used in patients with mild to moderate disease as in the current study, but it is unlikely that such a model of care (ie, with absence of F2F consultations) could or should be used in patients with more severe disease. In fact, this health care delivery model is intended to free up clinic space for exactly those patients who need more and intensive monitoring. Thirty-four percent of IBDsmart/IBDoc patients borrowed phones for the study, which may be important considering that people who borrow phones may not be adept at using smartphones in the first place.

It is possible that events outside the realm of HBI, SCCAI, or FC can happen to patients, such as pregnancy or reactions to immunosuppressing drugs. This is why it was important to have a comments box at the end of the IBDsmart entries to allow the reporting of these less common occurrences. Future iterations should consider adding questions about these less common occurrences so as to ensure that the gastroenterologist is aware of them in a timely manner.

Time taken to complete IBDsmart/IBDoc appointments was not formally timed, but the gastroenterologists reported that this was on average 2 minutes and no more than 5 minutes for each appointment when the patient was not flaring. When the physician was concerned about symptoms or elevated FC, more time was needed to arrange interventions such as phone call follow-up from IBD nurses or F2F clinical appointments, although this was not commonly required in this study and did not add time compared with the standard care group. In contrast, each standard F2F appointment took a minimum of 15 minutes to complete. Moreover, most physicians in the study reported that the apps reduced the amount of superfluous F2F care in the form of appointments with patients who were well. Future studies should ideally examine cost and time savings for patients and physicians more formally, but it is highly likely that

such reductions will be found. Finally, although IBDsmart and IBDoc are useful for concise care of IBD patients, it is important to be aware of the buy-in of the patients themselves. Of note is the European IMPACT survey indicated that more than half of patients surveyed felt that they were unable to inform their health care professional of something potentially important about their illness at least some of the time, and more still felt that their gastroenterologist should have asked more inquisitive questions to get a better understanding of their disease state.[29] Moreover, a Greek survey of patients reported that many patients want to receive more information from their doctor and that two-thirds wish for more IBD outpatient clinics.[30] Nevertheless, good usability and acceptability were reported by the patients in the present study, and many reported that the smartphone app–based care "made sense" to them in light of the present and future scarcity of specialist care for IBD patients. Moreover, some patients anecdotally reported that not having to come to see their gastroenterologist when they were in remission was very convenient in terms of reducing needless travel and time off work, although this feedback was not formally collected. Overall, smartphone app care should not be used on its own for initial appointments (ie, new diagnoses), severe IBD cases, or in patients who are unwilling, and some F2F follow-up may still be appropriate for many patients, albeit on a less frequent basis than is needed without the apps.

CONCLUSIONS

IBDsmart and IBDoc have been demonstrated to be acceptable, usable, and noninferior. This is encouraging in light of the need for more efficient use of specialist gastroenterologists' clinic time. However, gastroenterologists often felt that some information was not conveyed via the apps that would have been conveyed in an F2F setting, so future iterations should attempt to bring these questions into IBDsmart. Smartphone app care is most likely to be useful in willing patients with mild to moderate disease who are not brand new diagnoses. Future studies should directly compare a simple platform such as IBDsmart or IBDoc with a more complicated platform with psychoeducational components and the like.

SUPPLEMENTARY DATA

Supplementary data are available at *Inflammatory Bowel Diseases* online.

ACKNOWLEDGMENTS

BÜHLMANN Laboratories provided the fecal calprotectin kits and the IBDoc app for this study. CodeFlügel made the IBDsmart app at a cost. The following gastroenterologists were recruited for the study but are not on the author list: Dr. Richard Stein, Dr. Steven Burmeister, Dr. Ratna Pandey, Dr. Paul Frankish, Dr. Michael Burt, Dr. James Falvey, Professor Richard Gearry, Dr. Teresa Chalmers-Watson, and Dr. Ralf Lubcke. Kirsten Rosser was an IBD Nurse who recruited for the study but is not on the author list. Selina Brown is a Research Nurse who recruited for the study but is not on the author list.

REFERENCES

1. Bouma G, Strober W. The immunological and genetic basis of inflammatory bowel disease. *Nat Rev Immunol.* 2003;3:521–533.
2. Hjortswang H, Järnerot G, Curman B, et al. The influence of demographic and disease-related factors on health-related quality of life in patients with ulcerative colitis. *Eur J Gastroenterol Hepatol.* 2003;15:1011–1020.
3. Bernklev T, Jahnsen J, Lygren I, et al. Health-related quality of life in patients with inflammatory bowel disease measured with the Short Form-36: psychometric assessments and a comparison with general population norms. *Inflamm Bowel Dis.* 2005;11:909–918.
4. Coppell KJ, Galts CP, Huizing FY, et al. Annual incidence and phenotypic presentation of IBD in Southern New Zealand: an 18-year epidemiological analysis. *Inflamm Intest Dis.* 2018;3:32–39.
5. Molodecky NA, Soon IS, Rabi DM, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology.* 2012;142:46–54.e42; quiz e30.
6. Saibil F, Lai E, Hayward A, et al. Self-management for people with inflammatory bowel disease. *Can J Gastroenterol.* 2008;22:281–287.
7. Kennedy AP, Nelson E, Reeves D, et al. A randomised controlled trial to assess the effectiveness and cost of a patient orientated self management approach to chronic inflammatory bowel disease. *Gut.* 2004;53:1639–1645.
8. de Jong MJ, van der Meulen-de Jong AE, Romberg-Camps MJ, et al. Telemedicine for management of inflammatory bowel disease (myIBDcoach): a pragmatic, multicentre, randomised controlled trial. *Lancet.* 2017;390:959–968.
9. McCombie A, Langlotz T, Barclay M, et al. Systematic review of IBD smartphone apps for symptom monitoring and communication with healthcare professionals. *J mHealth.* 2016;3:36–41.
10. Con D, De Cruz P. Mobile phone apps for inflammatory bowel disease self-management: a systematic assessment of content and tools. *JMIR mHealth uHealth.* 2016;4:e13.
11. Spiegel B. Mobile health apps: e-wave of the future 2014. https://m.gastro.org/journals-publications/aga-perspectives/octobernovember2014/mobile-health-apps-ewave-of-the-future. Accessed April 13, 2015.
12. Harvey RF, Bradshaw JM. A simple index of Crohn's disease activity. *Lancet.* 1980;315:514.
13. Walmsley RS, Ayres RC, Pounder RE, et al. A simple clinical colitis activity index. *Gut.* 1998;43:29–32.
14. Regenbrecht H, Langlotz T, Ho C, et al. Field test of a questionnaire-based mobile health reporting system. In: Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design; Sydney, NSW, Australia: ACM; 2014:384–387.
15. Weber J, Ueberschlag M, Prica M, et al. Validation of a smartphone-based patient monitoring system measuring calprotectin as the therapy follow-up marker. Paper presented at: Annual Conference of the European Crohn's and Colitis Organisation; 2015; Barcelona, Spain.
16. Reinhard C, Weber J, Niederberger C, et al. Usability study of a smartphone-based patient monitoring system measuring Calprotectin for therapy follow up. Paper presented at: 10th Annual Conference of the European Crohn's and Colitis Organisation; 18-21 February 2015; Barcelona, Spain.
17. Costa F, Mumolo MG, Ceccarelli L, et al. Calprotectin is a stronger predictive marker of relapse in ulcerative colitis than in Crohn's disease. *Gut.* 2005;54:364–368.
18. Irvine EJ, Feagan B, Rochon J, et al. Quality of life: a valid and reliable measure of therapeutic efficacy in the treatment of inflammatory bowel disease. Canadian Crohn's Relapse Prevention Trial Study Group. *Gastroenterology.* 1994;106:287–296.
19. Higgins PDR, Schwartz M, Mapili J, et al. Patient defined dichotomous end points for remission and clinical improvement in ulcerative colitis. *Gut.* 2005;54:782–788.
20. 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.
21. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland, AL, eds. *Usability Evaluation in Industry*. London: Taylor and Francis; 1996: 107-114.

22. IBM Corp. *IBM SPSS Statistics for Windows*. 25th ed. Armonk, NY: IBM Corp; 2017.
23. Richardson G, Sculpher M, Kennedy A, et al. Is self-care a cost-effective use of resources? Evidence from a randomized trial in inflammatory bowel disease. *J Health Serv Res Pol.* 2006;11:225–230.
24. Bennebroek Evertsz F, Hoeks CC, Nieuwkerk PT, et al. Development of the patient Harvey Bradshaw index and a comparison with a clinician-based Harvey Bradshaw index assessment of Crohn's disease activity. *J Clin Gastroenterol.* 2013;47:850–856.
25. Sprangers MAG, Nieuwkerk PT, Bennebroek Evertsz' F, et al. The Patient Simple Clinical Colitis Activity Index (P-SCCAI) can detect ulcerative colitis (UC) disease activity in remission: a comparison of the P-SCCAI with clinician-based SCCAI and biological markers. *J Crohns Colitis.* 2013;7:890–900.
26. Jackson B, De Cruz P. Algorithms to facilitate shared decision-making for the management of mild-to-moderate ulcerative colitis. *Expert Rev Gastroenterol Hepatol.* 2018;12:1079–1100.
27. Walsh A, Kormilitzin A, Hinds C, et al. Defining faecal calprotectin thresholds as a surrogate for endoscopic and histological disease activity in ulcerative colitis—a prospective analysis. *J Crohns Colitis.* 2019;13:424–430.
28. Maaser C, Sturm A, Vavricka SR, et al; European Crohn's and Colitis Organisation [ECCO] and the European Society of Gastrointestinal and Abdominal Radiology [ESGAR]. ECCO-ESGAR Guideline for Diagnostic Assessment in IBD Part 1: initial diagnosis, monitoring of known IBD, detection of complications. *J Crohns Colitis.* 2019;13:144–164.
29. Lönnfors S, Vermeire S, Greco M, et al. IBD and health-related quality of life – discovering the true impact. *J Crohns Colitis.* 2014;8:1281–1286.
30. Viazis N, Mantzaris G, Karmiris K, et al; Hellenic Foundation of Gastroenterology and Nutrition. Inflammatory bowel disease: Greek patients' perspective on quality of life, information on the disease, work productivity and family support. *Ann Gastroenterol.* 2013;26:52–58.