World Journal of

Gastroenterology

World Journal of Gastroenterology

Contents

Contents

Contents

ABOUT COVER

Editorial Board Member of *World Journal of Gastroenterology*, Adam Cheifetz, MD, Professor of Medicine, Harvard Medical School, Director, Center for Inflammatory Bowel Disease, Medical Director, Infusion Services, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 022115, United States. aheifet@bidmc.harvard.edu

AIMS AND SCOPE

The primary aim of *World Journal of Gastroenterology* (*WJG*, *World J Gastroenterol*) is to provide scholars and readers from various fields of gastroenterology and hepatology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online. *WJG* mainly publishes articles reporting research results and findings obtained in the field of gastroenterology and hepatology and covering a wide range of topics including gastroenterology, hepatology, gastrointestinal endoscopy, gastrointestinal surgery, gastrointestinal oncology, and pediatric gastroenterology.

INDEXING/ABSTRACTING

The *WJG* is now indexed in Current Contents®/Clinical Medicine, Science Citation Index Expanded (also known as SciSearch®), Journal Citation Reports®, Index Medicus, MEDLINE, PubMed, PubMed Central, and Scopus. The 2021 edition of Journal Citation Report® cites the 2020 impact factor (IF) for *WJG* as 5.742; Journal Citation Indicator: 0.79; IF without journal self cites: 5.590; 5-year IF: 5.044; Ranking: 28 among 92 journals in gastroenterology and hepatology; and Quartile category: Q2. The *WJG*'s CiteScore for 2020 is 6.9 and Scopus CiteScore rank 2020: Gastroenterology is 19/136.

RESPONSIBLE EDITORS FOR THIS ISSUE

Observational Study

Assessing disease activity using the pediatric Crohn's disease activity index: Can we use subjective or objective parameters alone?

Amy Grant, Trudy Lerer, Anne M Griffiths, JS Hyams, Anthony Otley

ORCID number: Amy Grant 0000-0003-3605-0934; Trudy Lerer 0000-0002-0495-3074; Anne M Griffiths 0000-0001-8623-4665; JS Hyams 0000-0002-9769-236X; Anthony Otley 0000-0002-7371-9302.

Author contributions: Grant A acts as guarantor of this article; all authors contributed to the conception of the study; Grant A analyzed the data and wrote the manuscript; all other authors provided critical insight into analysis, interpretation of results, and revision of the manuscript; all authors approved the final version of the article, including the authorship list.

Institutional review board statement: Each individual study received institutional ethics review at each study site. Local Institutional Review Board approval was obtained by the IWK Health Center (Approval Nos: 3899, 2631, 2163).

Informed consent statement: All study participants, or their legal guardian, provided informed written consent prior to study enrollment.

Conflict-of-interest statement: Grant A and Lerer T have no conflicts of interest to disclose. Otley A is on the Advisory Board

Amy Grant, Anthony Otley, Division of Pediatrics, IWK Health, Halifax, NS B3K6R8, Canada

Trudy Lerer, Division of Research, Connecticut Children's Medical Center, Hartford, CT 06106, United States

Anne M Griffiths, Division of Gastroenterology Hepatology and Nutrition, The Hospital for Sick Children, Toronto, ON M5G1X8, Canada

JS Hyams, Division of Gastroenterology, Connecticut Children's Medical Center, Hartford, CT 06106, United States

Anthony Otley, Division of Pediatrics, Dalhousie University, Halifax, NS B3H4C3, Canada

Corresponding author: Anthony Otley, FRCPC, MD, Doctor, Professor, Division of Pediatrics, Dalhousie University, 5850/5980 University Avenue Halifax, Halifax, NS B3H4C3, Canada. arotley@dal.ca

Abstract

BACKGROUND
The pediatric Crohn's disease activity index (PCDAI) is used as a standard tool to assess disease activity in clinical trials for pediatric Crohn's disease.

AIM
To examine which items on the PCDAI drive assessment of disease activity, and how subgroups of subjective and objective items reflect change in disease state over time.

METHODS
Selective raw data from three prospectively collected datasets were combined, including 703 children with full PCDAI data at baseline, at 3-mo (Q1, $n = 670$), and 1-year (Q4, $n = 474$). Change in individual PCDAI scores from baseline to Q1 and to Q4 were examined using the non-weighted PCDAI.

RESULTS
Abdominal pain, well-being, weight, and stooling had the highest change scores over time. Objective indicators including albumin, abdominal exam, and height velocity followed. Change scores for well-being and abdominal exam did not explain significant variance at Q1 but were significant predictors at Q4 ($P < 0.001$ and $P < 0.05$). Subjective and objective subgroups of items predicted less variance

(18% and 22%) on total PCDAI scores at Q1 and Q4 compared to the full PCDAI, or a composite scale (both 32%) containing significant predictors.

CONCLUSION
Although subjective items on the PCDAI change the most over time, the full PCDAI or a smaller composite of items including a combination of subjective and objective components classifies disease activity better than a subgroup of either subjective or objective items alone. Reliance on subjective or objective items as stand-alone proxies for disease activity measurement could result in misclassification of disease state.

Key Words: Crohn's disease; Pediatric Crohn's disease activity index; Patient reported outcome measurement; Disease activity; Clinical trials; Pediatric

Core Tip: The pediatric Crohn's disease activity index (PCDAI) is commonly used to assess disease activity in clinical trials. The PCDAI is a multi-item index incorporating subjective (*e.g.*, patient well-being) and objective (*e.g.*, laboratory tests) items. In response to a call from the Food and Drug Administration our team reexamined functioning of this index. Although subjective items on the PCDAI changed the most over time, the full PCDAI or a smaller composite of items that includes both subjective and objective components better classifies disease activity. Use of subjective or objective items on their own may result in misclassification of disease state.

INTRODUCTION

The pediatric Crohn's disease activity index (PCDAI) has been the standard tool to assess clinical disease activity and response to treatment in clinical trials of pediatric Crohn's disease (CD) for the past two decades[1-4]. The PCDAI was developed by a group of clinicians and focuses on: (1) Subjective reporting of the degree of abdominal pain, stool pattern, and general well-being; (2) Extra-intestinal manifestations, such as fever, arthritis, rash, and uveitis; (3) Physical examination findings including abdominal pain, perirectal disease, extraintestinal manifestations, weight and height; and (4) Laboratory data, including hematocrit, erythrocyte sedimentation rate (ESR), and serum albumin[1] . This index has high inter-rater reliability, and good construct validity with physician global assessments of disease activity[1]. Validity has been further demonstrated in several studies assessing its psychometric properties[5-9].

Over time, concerns over both feasibility and short-term responsiveness to change in clinical status have arisen, due to the inclusion of perianal examination, as well as height velocity, which are not expected to change in a short-term (typically 8-12-wk period for induction of remission) clinical trial. These feasibility concerns are problematic in use of the PCDAI as the completion rate in routine practice of the PCDAI is approximately half (48% *vs* 98%)[10] that of the pediatric ulcerative colitis activity index (PUCAI), which does not include laboratory values, physical examination, or growth[11]. In response, other versions of the PCDAI have been developed, which have demonstrated increased feasibility: the abbreviated PCDAI[10, 12,13], the Short PCDAI[10], and the modified PCDAI[14], but they have had lower face, construct and discriminant validity than the original PCDAI[5]. Subsequently, a mathematically weighted version (wPCDAI) that removed redundant and low feasibility items was established[5]. Despite multiple versions, there remain concerns around construct validity due to the mismatch between clinical symptoms relied upon in the wPCDAI and objective markers of inflammation as visualized endoscopically, or

reflected *via* C-reactive protein or fecal calprotectin[15-17].

Traditionally, endpoints chosen for clinical trials have been based on multi-item indices[18,19], such as the PCDAI, which can incorporate symptoms, signs, laboratory tests, and endoscopic measures[20]. The PCDAI is still in use as an endpoint in many pediatric CD trials despite the feasibility concerns of these multi-item indices. As such, the United States Food and Drug Administration (FDA) sponsored the Gastroenterology Regulatory Endpoints and Advancement of Therapeutics (GREAT) workshops to reassess their use[20]. The FDA is seeking to redefine endpoints used in clinical trials, based on the need for objective measures of disease state. Additionally, recent regulatory guidance mandates that symptoms should be scored directly from the patient without a proxy[21]. This guidance recommends the development and use of Patient Reported Outcomes (PROs)[22-25] to measure clinical trial end-points, alongside objective markers of disease (such as endoscopic tools)[26] and health related quality of life[27,28].

The PCDAI is a composite of subjective items derived from a clinical observer, as well as objective items (laboratory studies), and as such it does not meet criteria set forth by the FDA for a PRO[22]. Although recent work has repurposed pieces of the PCDAI (ESR, stooling) with fecal calprotectin and C-reactive protein to develop a non-invasive index of mucosal healing in pediatric CD[29], there is no validated measure of subjective disease state in this population. A PRO for pediatric CD disease activity is still in the early stages of development[30] and is years away from implementation and use as an endpoint in clinical trials. Meanwhile, adult inflammatory bowel disease (IBD) trials are currently moving away from multi-item indices as endpoints and moving towards FDA guidance involving use of a PRO to measure functional components of disease alongside an objective marker of inflammation[31]. As such, as part of the GREAT workshops, the academic and pharmaceutical communities were asked to investigate and redefine the drivers of remission on the PCDAI given the urgent need for a PRO in pediatric CD.

In response to the FDA call we have set out to evaluate whether subgroups of items on the PCDAI could serve as an indicator of objective disease activity, or subjective disease activity until a PRO is ready for use. To date, post-hoc analyses of the PCDAI have not evaluated subcomponents that may serve as interim proxies of subjective or objective components, nor have they evaluated how individual items change over time. In the current study we evaluated the functioning of items on the PCDAI to determine how well subgroups of subjective and objective PCDAI items can capture disease activity.

MATERIALS AND METHODS

Ethics

Appropriate ethics approvals were attained across all participating sites at the original time of data collection. No additional ethics approvals were required for the purpose of this study.

Participants

Three prospectively collected datasets were merged and used for analysis, as follows.

The Registry[8]: The pediatric inflammatory bowel disease (IBD) collaborative research group registry. Data from Canada and the United States were collected on children (< 16 years old) newly diagnosed with IBD across 22 sites between 2002 and 2014. For children with CD, PCDAI, as well as clinical and demographic characteristics, physician global assessment of disease activity, and therapies were recorded at time of diagnosis or at study enrolment, which sometimes differed by up to a few weeks (hereinafter referred to as baseline), 30 d after diagnosis, and then quarterly. Children were not selected for inclusion based on disease severity. In the current study, data from baseline, first quarter (Q1), and at 1-year post diagnosis (Q4) were extracted for 414 children with PCDAI scores available.

The REACH study[4]: Clinical trial of infliximab use in pediatric CD. This randomized controlled trial of infliximab use included 112 children 6-17 years old who had moderate to severe CD (PCDAI > 30) at baseline. This trial involved a 10-wk induction phase, followed by a 1-year maintenance phase. Data at baseline, 10 wk (considered Q1), and 1-year (Q4) were extracted for the current study.

The IMAgINE trial[32]: Clinical trial of adalimumab use in pediatric CD. In this randomized controlled trial, 192 patients between 6 and 17 years old, with moderate to severe CD (PCDAI > 30) received open-label adalimumab at baseline and at 2 wk

before being randomized in blinded fashion at week 4 to high or low dose alternate weekly adalimumab maintenance for an additional 48 wk. Patients were stratified prior to randomization by week 4 responder status and prior exposure to infliximab. Data for the current study were extracted at baseline, 12 wk (Q1) and 52 wk post-treatment initiation (Q4).

Measures

The original version of the 11-item PCDAI (*i.e.,* non-weighted) was used in this study [1], as this was in keeping with how data were collected, scored, and analyzed in each of the original studies. The PCDAI was used to assess disease activity, with the acknowledgement that there is no current gold standard of disease activity in pediatric CD assessment. Severity for each item is assigned a score of 0 (normal), 5 (mild abnormality), or 10 (severe abnormality) except for hematocrit and ESR which are scored as 0, 2.5, or 5. A minimum total index score is zero, and a maximum score is 100. Remission was defined as a total PCDAI score ≤ 10, congruent with the REACH and IMAgINE trials[4,32].

Data acquisition

Relevant raw data were provided from each of the studies and were merged for group analysis. Basic demographic characteristics (age, gender, disease duration), and PCDAI scores at the above-mentioned time points were included. Physician Global Assessment (REACH, Registry), Subject and Parent Global Assessment (REACH), as well as IMPACT, a 35-item health related quality of life self-report measure for paediatric IBD patients[33] (REACH, Registry) were also provided, but were not included in the current analysis. Only data relevant to the proposed analyses were provided, thus, for further information on disease location, or types of treatment, please see the original manuscripts for the REACH[4] and IMAgINE[32] trials.

Inclusion criteria

Only participants with complete PCDAI data at baseline (i.e. had a score for each item on the questionnaire) were included. Thus, sample sizes for clinical trial datasets differ from the original reports, which used different missing data rules (*e.g.*, last value carried forward) as appropriate to the methodology of each respective trial. Sample sizes differ when examining change in PCDAI scores over time, as not all participants had complete data at each respective timepoint. Data were determined to be missing at random (*i.e.* did not attend scheduled follow-up appointment, some clinical or laboratory data parameters not collected due to established PCDAI feasibility issues).

Statistical analysis

Differences in demographic characteristics and baseline scores for participants with and without complete PCDAI data at each timepoint were compared using Independent samples Mann-Whitney U tests. PCDAI change scores were calculated by subtracting the score at baseline from the score at each follow-up time point (either Q1 or Q4, as appropriate). Repeated measures general linear models were run, which examined the mean change from baseline to Q1 and from baseline to Q4 for individual PCDAI items. Fisher's least significant difference tests for pairwise comparisons were run within the general linear model to examine the rank order of change scores across items. Multivariate linear regressions were run to examine the relative contribution of individual PCDAI change scores (entered in weighted units) from baseline to total disease activity at Q1 and Q4. Unstandardized beta values are reported. Subsequent models were run to examine the relative performance of subjective and objective item change scores that significantly contributed to the overall score at Q1 and Q4 follow-up, as well as models looking at only those items that significantly contributed to the overall score at each follow-up time point.

Serial receiver operating characteristic curves (ROC) were run to determine the optimal total score cut-point that differentiates between remitters and non-remitters for the two subscales containing subjective items and objective items. The best cut-off score maximized sensitivity and specificity, corresponding to the upper left shoulder of the ROC curve. The cut-points were then used to examine how well subjective and objective scales alone would similarly classify patients as achieving remission (remitters) *vs* non-remitters compared to the full PCDAI scale score.

All analyses were carried out using SPSS V24.0 (Chicago, IL, United States).

RESULTS

A total of 703 children, with full PCDAI data at baseline, were included in the current study (Table 1). Patients enrolled in the Registry were all at diagnosis, of varying disease severity, and treated per physician dictate, not by protocol. Patients enrolled in the REACH and IMAGINE studies had moderate to severe disease activity at study entry, had a mean disease duration of approximately two years since diagnosis, and were treated by the clinical trial protocol. Basic demographic information, including age, disease duration, and mean PCDAI scores, are presented separately for each dataset as the Registry sample was significantly different than each of the trials with respect to age at baseline (younger), age at diagnosis (younger), and baseline PCDAI (lower) (all $P < 0.001$). The data were then merged and analyzed together for all subsequent analyses, as the goal with merging the data was to have a diverse and representative sample. Sample sizes differ for each follow-up time point as some participants had incomplete PCDAI data at those follow-ups. There were some differences between those who had full PCDAI data at Q1 ($n = 670$), as participants who did not have full Q1 data ($n = 39$) were somewhat older (13.5 ± 2.3 *vs* 12.5 ± 2.8 years, $P = 0.059$), but were significantly younger at diagnosis (10.2 ± 2.9 *vs* 11.5 ± 2.8 years, $P = 0.004$), and had higher baseline PCDAI scores (42.1 ± 9.6 *vs* 34.9 ± 13.7, $P < 0.001$). Participants who did not have full PCDAI data at Q4 ($n = 235$) were significantly younger (12.0 ± 2.7 *vs* 12.8 ± 2.8, $P < 0.001$) and had lower baseline PCDAI scores (32.3 ± 14.7 *vs* 36.89 ± 12.7, $P < 0.001$) compared to those with full PCDAI data at Q4 ($n = 474$). There were no differences in gender distribution between those who did and did not have complete data at each time point (both $P > 0.53$).

The mean absolute change of individual PCDAI items was examined over time for individuals who had complete PCDAI data from baseline to Q1 ($n = 670$), and from baseline to Q4 ($n = 474$) (Figure 1). Of note is that accurate comparisons of mean change cannot be made between hematocrit or ESR and the other PCDAI items as they are scored on 5-point *vs* the 10-point scale of the other PCDAI items. The mean change of individual items was similar at both Q1 and Q4. Abdominal pain, well-being, weight, and stooling had a larger degree of change from baseline to Q1 and Q4 in comparison to the other items ($P < 0.05$). Other objective indicators of disease activity followed, with albumin, abdominal exam, height velocity, and ESR being within the next four most highly ranked items showing similar degrees of absolute change over time to both follow-up time points.

Two linear regression analyses were run to examine the contribution of individual PCDAI change scores to disease activity at Q1 and Q4 (Table 2). At Q1, abdominal pain, stooling, ESR, albumin, weight, height velocity, and perirectal examination were all significant predictors of disease activity (all $P < 0.05$). Three additional models were run: (1) Subjective; (2) Objective; or (3) A composite scale of PCDAI items that were significant predictors in the full model. All the models were significant (all $P < 0.001$), however, the proportion of variance accounted for by the PCDAI questions included was highest for the total and composite scales (both $R^2 = 0.32$), where less variance was accounted for on the models including only the subjective ($R^2 = 0.18$) and objective ($R^2 = 0.21$) items. Similar results were found at Q4, except that the PCDAI change scores from baseline to Q4 were also significant for well-being ($P < 0.001$), abdominal exam ($P < 0.05$), and extraintestinal manifestations ($P < 0.05$). The three additional models also all accounted for a significant amount of variance (all $P < 0.001$), with the subjective scale at this time point also including well-being, and the objective scale included abdominal exam and extraintestinal manifestations, as previously discussed. Similar to Q1, more variance was accounted for by the full and composite models (both $R^2 = 0.45$), than for the subjective ($R^2 = 0.22$) or objective ($R^2 = 0.36$) models.

ROC curves were run to examine how well these sub-scales of subjective and objective items would function on their own, as a proxy to measure disease activity. Items included on the subjective and objective subscales were made based on which item change scores significantly predicted disease activity at the first follow-up (Q1) as this is a similar timeframe to many clinical trials and corresponds with the degree of clinical change observed during a feasible assessment time frame. Cut-points to define remission were defined at Q1 ($n = 670$), using a composite of the significant subjective items (abdominal pain and stooling) (Figure 2A) and objective items (ESR, albumin, weight, height velocity, and perirectal examination) (Figure 2B). A cut-point of ≤ 2.5 (out of a maximum score of 20) best defined remission for the subjective composite, based on an area under the curve of 0.84 (95%CI: 0.80-0.87), with a sensitivity of 74% and a specificity of 83%. For the objective composite, a cut-point of ≤ 3.75 (out of 45) defined remission, with an area under the curve of 0.85 (95%CI: 0.82-0.88), sensitivity of 71% and specificity of 78%.

Table 1 Descriptive statistics

	Registry (*n* = 414[1])		REACH (*n* = 101[1])		IMAgINE (*n* = 188[1])	
	Range	Median (IQR)	Range	Median (IQR)	Range	Median (IQR)
Age at baseline	1.5-16.1	12.2 (3.6)	6.0-17.8	13.8 (3.2)	6-17.0	14.0 (4.0)
Age at diagnosis	1.5-15.9	12.2 (3.6)	4.8-16.6	11.8 (4.1)	2.4-16.4	10.6 (3.2)
Gender (female), *n* (%)	180 (43.5)		40 (39.6)		82 (43.6)	
	Range	mean ± SD	Range	mean ± SD	Range	mean ± SD
PCDAI BL	0-82.5	31.39 ± 15.34	25-62.5	41.14 ± 8.18	7.5-62.5	40.97 ± 7.48
PCDAI Q1	0-52.5	10.25 ± 9.75	0-40	9.77 ± 8.55	0-57.5	16.78 ± 12.64
PCDAI Q4	0-50.0	8.58 ± 9.21	0-70	12.07 ± 14.10	0-70	19.56 ± 16.34

[1]The baseline characteristics describe participants with full pediatric Crohn's disease activity index (PCDAI) data (*i.e.* had individual PCDAI scores collected) at baseline. Q1 PCDAI scores describe participants with full baseline and Q1 PCDAI data available (*n* = 414, 97, and 159 participants for the Registry, REACH, and IMAgINE datasets respectively). Q4 PCDAI scores describe participants with full baseline and Q1 PCDAI data available (*n* = 216, 75, and 183 participants for the Registry, REACH, and IMAgINE datasets respectively). PCDAI: Pediatric Crohn's disease activity index.

Remission status at Q1 and Q4 was examined using subscales of subjective and objective items with the cut-points identified above to examine the comparability of using subscales compared to the full PCDAI (Table 3). This table demonstrates the proportion of patients that would be classified the same or different as the original PCDAI scale by using subscale scores only, thus either in remission ("remitters") or having active disease ("non-remitters"). At Q1, of those classified as in remission using the full PCDAI, only 73.8% would be classified as in remission using the subjective subscale, and thus 26.2% of patients would be classified incorrectly. Using the objective scale, 71.5% of patients deemed remitters using the full PCDAI would be classified as being in remission. For patients not in remission using the full PCDAI, 16.7% would be differently classified using the subjective scale, and 21.2% using the objective scale. At Q4 73.7% of those in remission according to the full PCDAI scale would be in remission based on the subjective subscale, and 74.5% would be in remission according to the objective subscale. For those with active disease at Q4, 77.5% would be similarly defined as having active disease using the subjective subscale, and 72.0% would have active disease using the objective subscale.

DISCUSSION

In response to the FDA call for reevaluation of subcomponents of the PCDAI, and through work generated through the GREAT-III workshops, we have evaluated items on the PCDAI to determine the suitability of these subjective components in assessing disease activity on their own–while we await the development of a PRO for pediatric CD that fulfills FDA guidance. The results of this study indicate that relying on these items, as they are currently measured (*i.e.*, clinician report, nominal variable scoring), will not suffice as a substitution for a robust PRO outcome measure. However, the results do indicate that subjective items such as abdominal pain and stooling in the short-term, and well-being in a longer-term measurement are adequate indicators of disease activity and response to treatment.

We examined both absolute change in scores over time, in addition to looking at what change scores explained unique variance in disease activity at a 3-mo and 1-year follow-up. Items that changed the most over this period were similar. Subjective items including abdominal pain, well-being, and stooling were consistently among the top four items that changed the most over these periods. The rank order of the remaining items varied across these two time points, although objective markers of disease activity including weight, albumin, and abdominal exam were amongst the next group of items that changed the most over time. Results were similar when looking at the unique contributions of the change in item scores to the follow-up total PCDAI score at each time point. However, well-being was not a unique predictor at the 3-mo follow-up, nor was abdominal exam. At the 1-year follow-up however, almost all the change scores from baseline to 1-year added unique variance to the total score, except for hematocrit.

Table 2 Linear regression analysis examining the contribution of pediatric Crohn's disease activity index change scores to Q1 and Q4 pediatric Crohn's disease activity index disease activity

	PCDAI∆ Q1 to BL (*n* = 670)			
	Full scale	**Subjective**	**Objective**	**Composite**
	β values	**β values**	**β values**	**β values**
1. Abdominal pain	0.57[b]	0.80[b]		0.65[b]
2. Stooling	0.42[b]	0.53[b]		0.45[b]
3. Well-being	0.13			
4. HCT	-0.09			
5. ESR	0.94[b]		1.27[b]	0.96[b]
6. Albumin	0.46[b]		0.41[b]	0.45[b]
7. Weight	0.30[b]		0.52[b]	0.32[b]
8. Height Velocity	0.50[b]		0.56[b]	0.48[b]
9. Abdominal exam	0.08			
10. Perirectal exam	0.47[b]		0.45[a]	0.48[b]
11. Extraintestinal	0.13			
R^2	0.32	0.18	0.21	0.32
	PCDAI∆ Q4 to BL (*n* = 474)			
1. Abdominal Pain	0.45[b]	0.65[b]		0.45[b]
2. Stooling	0.42[a]	0.52[b]		0.42[a]
3. Well-being	0.58[b]	0.87[b]		0.58[b]
4. HCT	-0.01			
5. ESR	1.32[b]		1.56[b]	1.32[b]
6. Albumin	0.81[b]		0.84[b]	0.81[b]
7. Weight	0.49[b]		0.63[b]	0.49[b]
8. Height velocity	0.45[b]		0.48[b]	0.45[b]
9. Abdominal exam	0.35[a]		0.83[b]	0.35[a]
10. Perirectal exam	0.78[b]		0.90[b]	0.78[b]
11. Extraintestinal	0.42[a]		0.67[a]	0.42[b]
R^2	0.45	0.22	0.36	0.45

[a]$P < 0.05$.

[b]$P \leq 0.001$. PCDAI: Pediatric Crohn's disease activity index; HCT: Hematopoietic cell transplantation; ESR: Erythrocyte sedimentation rate.

Dominant PCDAI items (*i.e.,* items that contributed significant, unique variance) were similar to previous work examining functioning of the PCDAI and the contribution of individual items to change in disease activity[9]. At the 3-mo follow-up, well-being, hematocrit, abdominal exam, and extraintestinal manifestations did not explain unique variance in the total PCDAI score. This is similar to previous work demonstrating the limited responsiveness of specific items on the PCDAI, where abdominal exam, height, and hematocrit were removed from the currently recommended weighted PCDAI as they were found to be redundant in predicting change in disease activity as measured by physician global assessment[5]. At the 1-year follow-up, change scores for all items contributed significant unique variance to the total PCDAI score except for hematocrit. The strength in the current work is the examination of change scores over a 1-year period, highlighting the composite nature of the PCDAI, and that the change in items such as height velocity, abdominal exam, and extraintestinal manifestations take a longer time to demonstrate statistically unique or meaningful change but are nonetheless important components of overall well-being and functioning in patients with CD.

Table 3 Correct and incorrect identification of remitters and non-remitters

| Full PCDAI-remission status	Subjective-remission status		Objective-remission status	
	Remitters	Non-remitters	Remitters	Non-remitters
	n (%)	*n* (%)	*n* (%)	*n* (%)
Remitters at Q1 (*n* = 397)	292 (73.8)	105 (26.2)	284 (71.5)	113 (28.5)
Non-remitters at Q1 (*n* = 273)	46 (16.7)	227 (83.3)	58 (21.2)	215 (78.8)
Remitters at Q4 (*n* = 274)	202 (73.7)	72 (26.3)	204 (74.5)	70 (25.3)
Non-remitters at Q4 (*n* = 200)	45 (22.5)	155 (77.5)	56 (28.0)	144 (72.0)

Subjective subscale is composed of abdominal pain and stooling; Objective subscale is composed of erythrocyte sedimentation rate, albumin, weight, height velocity, and perirectal examination. Shaded cells = number of patients that would be incorrectly (or differently) classified as having active or inactive disease based on the subjective components alone (*e.g.*, At the Q1 time point, 397 patients were classified as being in remission, but based on the subjective only scoring, 292 (73.8%) of those patients would be classified as being in remission, while the other 26% would be classified as having active disease). PCDAI: Pediatric Crohn's disease activity index.
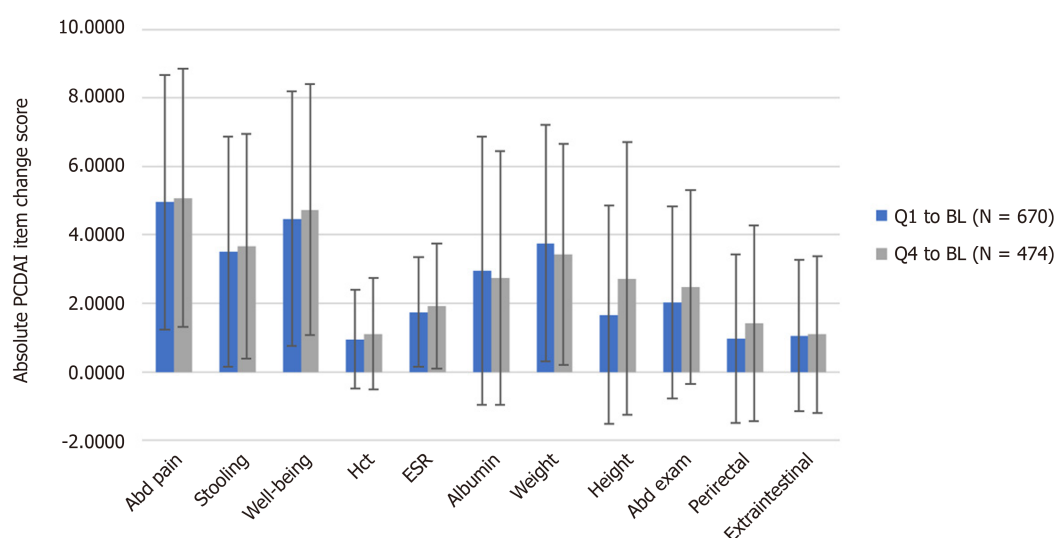
Figure 1 Mean change over time (from baseline to Q1 and Baseline to Q4). Hematocrit and erythrocyte sedimentation rate are scored on a 0-5-point scale, whereas all other items are scored on a 0-10 point scale.

Further examination of the subjective and objective subcomponents highlighted that change in both subcomponents over time does contribute significantly to disease activity. However, change in either subjective or objective groups of items on their own did not perform as well as the full-scale, or to a subgroup of questions (items that were significant unique predictors of total scores) that served as a composite measure of disease activity. If these subcomponents were used to classify patients into remission or active disease, a large portion of patients would be classified differently compared to using the full-scale. The outcomes at the 3-mo and 1-year assessments were similar, where approximately 25% of patients in remission with the full scale would be classified as having active disease with either the subjective or objective scales.

For patients with active disease, the percentage of patients misclassified ranged from about 16% to 22% with the subjective scale, whereas the range was 21% to 28% misclassified with the objective scale at both follow-up time points. Although the composite of items showed similar performance to the full PCDAI, the goal of this work was not to recommend or endorse a composite PCDAI measure as this work has already been completed[5,10], and our findings produced similar results as previously described. Rather, the focus of the current work around these smaller subgroups of items reflects the limited ability of either subjective or objective components alone, as measured on the PCDAI, to fully characterize disease activity.
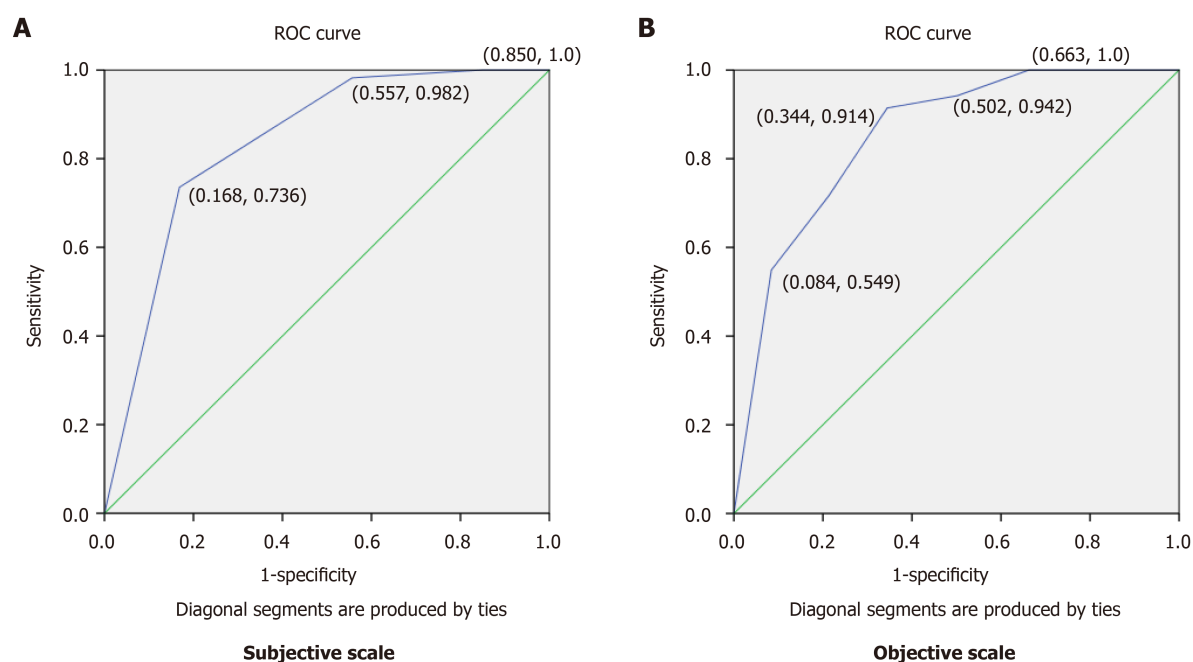
Figure 2 Receiver operator characteristic curve to identify the equivalent remission cut point of the subjective and objective scales. A: Subjective scale; B: Objective scale. ROC: Receiver operator characteristic.

Our study has several limitations, where the main limitation in the current work is the lack of "gold standard" measurement of disease activity beyond physician global assessment. The examination of the change in subjective and objective items on the PCDAI was compared over time but in relation to measurement of disease activity with the same measure at follow-up time points. Other work in trying to assess the psychometric properties of existing measures of disease activity in pediatric CD has found limited correlation with objective measurement of mucosal inflammation with the Simple Endoscopic Score for CD[15,16]. Additionally, work to identify biomarkers that accurately identify disease activity in pediatric CD is ongoing[34], but currently suggests that fecal calprotectin and C-reactive protein are reliable markers with utility in management of a patient's condition[35]. These indices also demonstrate limited correlation with C-reactive protein, and a poor correlation with fecal calprotectin[15, 16]. Although there were some differences (younger at diagnosis, higher disease activity at baseline) between participants who had complete follow-up data at 1 year and those who did not, these differences were not considered problematic in the context of the current study. The objective of this study was to examine performance of the PCDAI rather than response to treatment, thus differences between those with and without complete data were not thought to be influential factors that would affect outcomes. In the absence of a universal gold standard measure, the current study is a first step in further understanding the functioning of subjective and objective assessment of disease activity as measured on the primary outcome measure used in clinical trials-the PCDAI.

Without a pediatric CD PRO ready for use, it is tempting to consider using measures already in existence to serve as an interim substitution. However, our study found that looking at subjective items alone in a large sample of patients with a range of disease activity is likely to misclassify many of these patients by relying on subjective items alone. Although subjective items on the PCDAI change the most over time, these items do not capture disease activity as well as a composite of PCDAI items containing both subjective and objective components, or in comparison to the full PCDAI scale. It is not surprising that a composite scale might outperform individual items given the protean nature of CD. This may be particularly relevant when looking over longer periods of time when changes in growth parameters may be more evident.

CONCLUSION

Our analysis highlights the limitations of relying only on subjective or objective items as currently measured on the PCDAI as stand-alone proxies for disease activity

measurement in pediatric patients with CD. Measuring improvements in health outcomes should be informed by evidence-based information produced by patients [36]. Where improvement in disease state may not reflect improvements in a patients well-being[37], our data reiterate that complementary subjective (*i.e.,* PROs)[28] and objective markers of disease[26] should be used to evaluate the outcomes of treatment.

ARTICLE HIGHLIGHTS

Research background
The pediatric Crohn's disease activity index (PCDAI) is a standard tool to assess disease activity in clinical trials for pediatric Crohn's disease. Over time, concerns over both feasibility and short-term responsiveness to change in clinical status have arisen. Based on feasibility concerns, and new guidance recommending that symptoms are scored directly from a patient, the PCDAI was reexamined.

Research motivation
In response to a call from the Food and Drug Administration our team reexamined functioning of this index.

Research objectives
The objective of this study was to examine which items on the PCDAI drive assessment of disease activity, and how subgroups of subjective and objective items reflect change in disease state over time.

Research methods
We retrospectively examined data from three completed studies – one registry study and two clinical trials involving pediatric patients with Crohn's disease. Data was collected at baseline, at 3-mo (Q1) and 1-year (Q4). Change in individual PCDAI scores from baseline to Q1 and to Q4 were examined using the non-weighted PCDAI.

Research results
Abdominal pain, well-being, weight, and stooling had the highest change scores over time. Objective markers of disease activity including weight, albumin, and abdominal exam were amongst the next group of items that changed the most over time. Subjective and objective subgroups of items predicted less variance on total PCDAI scores at Q1 and Q4 compared to the full PCDAI, or a composite scale containing significant predictors.

Research conclusions
Although subjective items on the PCDAI change the most over time, the full PCDAI or a smaller composite of items including a combination of subjective and objective components classifies disease activity better than a subgroup of either subjective or objective items alone. However, the results do indicate that subjective items such as abdominal pain and stooling in the short-term, and well-being in a longer-term measurement are adequate indicators of disease activity and response to treatment. Reliance on subjective or objective items as stand-alone proxies for disease activity measurement could result in misclassification of disease state.

Research perspectives
Subjective or objective items alone, as currently measured on the PCDAI, do not serve as a substitute for a robust patient-reported outcome measure. Complementary subjective (*i.e.,* patient reported outcome measures) and objective markers of disease should be used to evaluate the outcomes of treatment.

REFERENCES

1   **Hyams JS**, Ferry GD, Mandel FS, Gryboski JD, Kibort PM, Kirschner BS, Griffiths AM, Katz AJ, Grand RJ, Boyle JT. Development and validation of a pediatric Crohn's disease activity index. *J Pediatr Gastroenterol Nutr* 1991; **12**: 439-447 [PMID: 1678008 DOI: 10.1097/00005176-199105000-00005]

2   **Johnson T**, Macdonald S, Hill SM, Thomas A, Murphy MS. Treatment of active Crohn's disease in children using partial enteral nutrition with liquid formula: a randomised controlled trial. *Gut* 2006; **55**: 356-361 [PMID: 16162683 DOI: 10.1136/gut.2004.062554]

3   **Pappa HM**, Mitchell PD, Jiang H, Kassiff S, Filip-Dhima R, DiFabio D, Quinn N, Lawton RC, Varvaris M, Van Straaten S, Gordon CM. Treatment of vitamin D insufficiency in children and adolescents with inflammatory bowel disease: a randomized clinical trial comparing three regimens. *J Clin Endocrinol Metab* 2012; **97**: 2134-2142 [PMID: 22456619 DOI: 10.1210/jc.2011-3182]

4   **Hyams J**, Crandall W, Kugathasan S, Griffiths A, Olson A, Johanns J, Liu G, Travers S, Heuschkel R, Markowitz J, Cohen S, Winter H, Veereman-Wauters G, Ferry G, Baldassano R; REACH Study Group. Induction and maintenance infliximab therapy for the treatment of moderate-to-severe Crohn's disease in children. *Gastroenterology* 2007; **132**: 863-73; quiz 1165 [PMID: 17324398 DOI: 10.1053/j.gastro.2006.12.003]

5   **Turner D**, Griffiths AM, Walters TD, Seah T, Markowitz J, Pfefferkorn M, Keljo D, Waxman J, Otley A, LeLeiko NS, Mack D, Hyams J, Levine A. Mathematical weighting of the pediatric Crohn's disease activity index (PCDAI) and comparison with its other short versions. *Inflamm Bowel Dis* 2012; **18**: 55-62 [PMID: 21351206 DOI: 10.1002/ibd.21649]

6   **Kundhal PS**, Critch JN, Zachos M, Otley AR, Stephens D, Griffiths AM. Pediatric Crohn Disease Activity Index: responsive to short-term change. *J Pediatr Gastroenterol Nutr* 2003; **36**: 83-89 [PMID: 12500001 DOI: 10.1097/00005176-200301000-00016]

7   **Otley A**, Loonen H, Parekh N, Corey M, Sherman PM, Griffiths AM. Assessing activity of pediatric Crohn's disease: which index to use? *Gastroenterology* 1999; **116**: 527-531 [PMID: 10029610 DOI: 10.1016/s0016-5085(99)70173-3]

8   **Hyams J**, Markowitz J, Otley A, Rosh J, Mack D, Bousvaros A, Kugathasan S, Pfefferkorn M, Tolia V, Evans J, Treem W, Wyllie R, Rothbaum R, del Rosario J, Katz A, Mezoff A, Oliva-Hemker M, Lerer T, Griffiths A; Pediatric Inflammatory Bowel Disease Collaborative Research Group. Evaluation of the pediatric crohn disease activity index: a prospective multicenter experience. *J Pediatr Gastroenterol Nutr* 2005; **41**: 416-421 [PMID: 16205508 DOI: 10.1097/01.mpg.0000183350.46795.42]

9   **Turner D**, Griffiths AM, Walters TD, Seah T, Markowitz J, Pfefferkorn M, Keljo D, Otley A, Leleiko NS, Mack D, Hyams J, Levine A. Appraisal of the pediatric Crohn's disease activity index on four prospectively collected datasets: recommended cutoff values and clinimetric properties. *Am J Gastroenterol* 2010; **105**: 2085-2092 [PMID: 20372111 DOI: 10.1038/ajg.2010.143]

10   **Kappelman MD**, Crandall WV, Colletti RB, Goudie A, Leibowitz IH, Duffy L, Milov DE, Kim SC, Schoen BT, Patel AS, Grunow J, Larry E, Fairbrother G, Margolis P. Short pediatric Crohn's disease activity index for quality improvement and observational research. *Inflamm Bowel Dis* 2011; **17**: 112-117 [PMID: 20812330 DOI: 10.1002/ibd.21452]

11   **Turner D**, Hyams J, Markowitz J, Lerer T, Mack DR, Evans J, Pfefferkorn M, Rosh J, Kay M, Crandall W, Keljo D, Otley AR, Kugathasan S, Carvalho R, Oliva-Hemker M, Langton C, Mamula P, Bousvaros A, LeLeiko N, Griffiths AM; Pediatric IBD Collaborative Research Group. Appraisal of the pediatric ulcerative colitis activity index (PUCAI). *Inflamm Bowel Dis* 2009; **15**: 1218-1223 [PMID: 19161178 DOI: 10.1002/ibd.20867]

12   **Loonen HJ**, Griffiths AM, Merkus MP, Derkx HH. A critical assessment of items on the Pediatric Crohn's Disease Activity Index. *J Pediatr Gastroenterol Nutr* 2003; **36**: 90-95 [PMID: 12500002 DOI: 10.1097/00005176-200301000-00017]

13   **Shepanski MA**, Markowitz JE, Mamula P, Hurd LB, Baldassano RN. Is an abbreviated Pediatric Crohn's Disease Activity Index better than the original? *J Pediatr Gastroenterol Nutr* 2004; **39**: 68-72 [PMID: 15187784 DOI: 10.1097/00005176-200407000-00014]

14   **Leach ST**, Nahidi L, Tilakaratne S, Day AS, Lemberg DA. Development and assessment of a modified Pediatric Crohn Disease Activity Index. *J Pediatr Gastroenterol Nutr* 2010; **51**: 232-236 [PMID: 20479686 DOI: 10.1097/MPG.0b013e3181d13609]

15   **Turner D**, Levine A, Walters TD, Focht G, Otley A, López VN, Koletzko S, Baldassano R, Mack D, Hyams J, Griffiths AM. Which PCDAI Version Best Reflects Intestinal Inflammation in Pediatric Crohn Disease? *J Pediatr Gastroenterol Nutr* 2017; **64**: 254-260 [PMID: 27050050 DOI: 10.1097/MPG.0000000000001227]

16   **Zubin G**, Peter L. Predicting Endoscopic Crohn's Disease Activity Before and After Induction Therapy in Children: A Comprehensive Assessment of PCDAI, CRP, and Fecal Calprotectin. *Inflamm Bowel Dis* 2015; **21**: 1386-1391 [PMID: 25851564 DOI: 10.1097/MIB.0000000000000388]

17   **Yu Y**, Zhao H, Luo Y, Lou J, Chen J, Fang Y. Poor Concordance Between Clinical Activity and Endoscopic Severity in Pediatric Crohn's Disease: Before and After Induction Therapy. *Dig Dis Sci* 2021 [PMID: 33818661 DOI: 10.1007/s10620-021-06917-3]

18   **Brody T**. Clinical trials: study design, endpoints and biomarkers, drug safety, and FDA and ICH guidelines. 2ne ed. Academic press, 2016

19   **Green DJ**, Burnham JM, Schuette P, Liu XI, Maas BM, Yao L, McCune SK, Chen J, van den Anker

JN, Burckart GJ. Primary Endpoints in Pediatric Efficacy Trials Submitted to the US FDA. *J Clin Pharmacol* 2018; **58**: 885-890 [PMID: 29663424 DOI: 10.1002/jcph.1109]

20   **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 [PMID: 25127678 DOI: 10.1053/j.gastro.2014.08.003]

21   **Williet N**, Sandborn WJ, Peyrin-Biroulet L. Patient-reported outcomes as primary end points in clinical trials of inflammatory bowel disease. *Clin Gastroenterol Hepatol* 2014; **12**: 1246-56. e6 [PMID: 24534550 DOI: 10.1016/j.cgh.2014.02.016]

22   **U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research.** U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research; U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006; **4**: 79 [PMID: 17034633 DOI: 10.1186/1477-7525-4-79]

23   **Black N**. Patient reported outcome measures could help transform healthcare. *BMJ* 2013; **346**: f167 [PMID: 23358487 DOI: 10.1136/bmj.f167]

24   **Weldring T**, Smith SM. Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). *Health Serv Insights* 2013; **6**: 61-68 [PMID: 25114561 DOI: 10.4137/HSI.S11093]

25   **McLeod LD**, Coon CD, Martin SA, Fehnel SE, Hays RD. Interpreting patient-reported outcome results: US FDA guidance and emerging methods. *Expert Rev Pharmacoecon Outcomes Res* 2011; **11**: 163-169 [PMID: 21476818 DOI: 10.1586/erp.11.12]

26   **Ghosh S**. Are patient-reported outcome measures the way to go in inflammatory bowel disease? *Can J Gastroenterol Hepatol* 2014; **28**: 535 [PMID: 25390614 DOI: 10.1155/2014/748938]

27   **Matza LS**, Patrick DL, Riley AW, Alexander JJ, Rajmil L, Pleil AM, Bullinger M. Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. *Value Health* 2013; **16**: 461-479 [PMID: 23796280 DOI: 10.1016/j.jval.2013.04.004]

28   **Ruemmele FM**, Hyams JS, Otley A, Griffiths A, Kolho KL, Dias JA, Levine A, Escher JC, Taminiau J, Veres G, Colombel JF, Vermeire S, Wilson DC, Turner D. Outcome measures for clinical trials in paediatric IBD: an evidence-based, expert-driven practical statement paper of the paediatric ECCO committee. *Gut* 2015; **64**: 438-446 [PMID: 24821616 DOI: 10.1136/gutjnl-2014-307008]

29   **Cozijnsen MA**, Ben Shoham A, Kang B, Choe BH, Choe YH, Jongsma MME, Russell RK, Ruemmele FM, Escher JC, de Ridder L, Koletzko S, Martín-de-Carpi J, Hyams J, Walters T, Griffiths A, Turner D. Development and Validation of the Mucosal Inflammation Noninvasive Index For Pediatric Crohn's Disease. *Clin Gastroenterol Hepatol* 2020; **18**: 133-140. e1 [PMID: 30981008 DOI: 10.1016/j.cgh.2019.04.012]

30   **Crane MB**, Grant A, Hyams J, Croft N, Mack DR, Hussey S, Sylvester F, Huynh HQ, Turner D, Otley A. A132 Development of the TUMMY-CD, A symptoms-based disease activity patient reported outcome (PRO) for pediatric Crohn's disease. *J Can Assoc Gastroenterol* 2018; **1**: 197-198 [DOI: 10.1093/jcag/gwy009.132]

31   **Bojic D**, Bodger K, Travis S. Patient Reported Outcome Measures (PROMs) in Inflammatory Bowel Disease: New Data. *J Crohns Colitis* 2017; **11**: S576-S585 [PMID: 27797917 DOI: 10.1093/ecco-jcc/jjw187]

32   **Hyams JS**, Griffiths A, Markowitz J, Baldassano RN, Faubion WA Jr, Colletti RB, Dubinsky M, Kierkus J, Rosh J, Wang Y, Huang B, Bittle B, Marshall M, Lazar A. Safety and efficacy of adalimumab for moderate to severe Crohn's disease in children. *Gastroenterology* 2012; **143**: 365-74. e2 [PMID: 22562021 DOI: 10.1053/j.gastro.2012.04.046]

33   **Otley A**, Smith C, Nicholas D, Munk M, Avolio J, Sherman PM, Griffiths AM. The IMPACT questionnaire: a valid measure of health-related quality of life in pediatric inflammatory bowel disease. *J Pediatr Gastroenterol Nutr* 2002; **35**: 557-563 [PMID: 12394384 DOI: 10.1097/00005176-200210000-00018]

34   **Dulai PS**, Peyrin-Biroulet L, Danese S, Sands BE, Dignass A, Turner D, Mantzaris G, Schölmerich J, Mary JY, Reinisch W, Sandborn WJ. Approaches to Integrating Biomarkers Into Clinical Trials and Care Pathways as Targets for the Treatment of Inflammatory Bowel Diseases. *Gastroenterology* 2019; **157**: 1032-1043. e1 [PMID: 31228441 DOI: 10.1053/j.gastro.2019.06.018]

35   **Dragoni G**, Innocenti T, Galli A. Biomarkers of Inflammation in Inflammatory Bowel Disease: How Long before Abandoning Single-Marker Approaches? *Dig Dis* 2021; **39**: 190-203 [PMID: 32942275 DOI: 10.1159/000511641]

36   **Gabriel SE**, Normand SL. Getting the methods right--the foundation of patient-centered outcomes research. *N Engl J Med* 2012; **367**: 787-790 [PMID: 22830434 DOI: 10.1056/NEJMp1207437]

37   **Breitscheidel L**, Stamenitis S. Using patient-reported outcome assessments in clinical practice and their importance in risk management. *J Med Econ* 2009; **12**: 180-181 [PMID: 19691445 DOI: 10.3111/13696990903216278]