Development and Validation of the IBD-REFER Criteria: Early Referral for Suspected Inflammatory Bowel Diseases in Adults and Children

Ohad Atia, MD,[*] Adi Shosberger, MBA,[*] Gili Focht, BSc, MBA,[*] Oren Ledder, MD,[*] Raffi Lev-Tzion, MD,[*] Dan Navon, MBA,[*] Amit Assa, MD,[†] Baruch Yerushalmi, MD,[‡] Ron Shaoul, MD,[§] Dror S. Shouval, MD,[¶] Ariella Bar-Gil Shitrit, MD,[‖] Benjamin Koslowsky, MD,[‖] Iris Dotan, MD, PhD,[**] Revital Kariv, MD,[††] Eitan Lavon, MD,[††] and Dan Turner, MD, PhD[*]*

Background: Early treatment of inflammatory bowel disease (IBD) is associated with positive outcomes but a significant diagnostic delay has been reported in most countries.

Aim: We aimed to develop and validate IBD-REFER criteria, intended for primary care physicians, to screen patients at risk for IBD.

Methods: A Delphi group of 10 experts generated a list of symptoms associated with the onset of IBD, supplemented by a review of the literature. The list was reduced in an iterative process and graded based on importance. For data-driven statistical formatting, the charts of 200 IBD (100 children, 100 adults) and 100 non-IBD controls but with gastrointestinal symptoms were reviewed. The IBD-REFER items were scored for each subject, as well as the contending Red Flag criteria from the International Organization for the Study of IBD. External validation was performed on additionally enrolled cohorts of 100 IBD patients and 50 controls.

Results: The Delphi process retained 5 items as major criteria (≥1 item required for early referral) and 11 as minor (≥2 items required). Following the removal of uninformative items and further formatting in the data-driven stage, 10 core items were retained: 3 as major and 7 as minor. In the external validation, the final IBD-REFER criteria had a sensitivity/specificity of 98%/96% in adults and 96%/96% in children, significantly higher than achieved by the Red Flag criteria (71%/84% and 60%/88%, respectively; $P < 0.001$).

Conclusion: The IBD-REFER criteria may guide the selection of patients for expedited gastrointestinal investigation.

Lay Summary: Early treatment of inflammatory bowel disease (IBD) is associated with positive outcomes but a significant diagnostic delay has been reported in most countries. Therefore, we developed and validated IBD-REFER criteria, intended for primary care physicians, to screen patients at risk for IBD.

Key Words: inflammatory bowel diseases, diagnostic delay, screening criteria

INTRODUCTION

The prevalence of inflammatory bowel disease (IBD) is constantly rising in most countries with 15%–20% of cases presenting during childhood.[1, 2] IBD is often characterized by nonspecific symptoms and therefore the presentation may be nonspecific, particularly in Crohn's disease (CD). It may overlap with other gastrointestinal conditions, such as irritable bowel syndrome (IBS), celiac disease, lactose intolerance, and dyspepsia, or may present solely with extraintestinal manifestations.[3, 4] Selecting those who require prompt referral to a gastroenterologist may thus be challenging and lead to a significant diagnostic delay.[5, 6] Because early treatment in IBD is associated with improved disease outcome,[7] there is a need to develop effective strategies for guiding early referral of those with signs and symptoms suggestive of IBD.

Received for publications January 17, 2020; Editorial Decision February 28, 2020.

*Juliet Keidan Institute of Pediatric Gastroenterology Hepatology and Nutrition, Shaare Zedek Medical Center, The Hebrew University of Jerusalem, Jerusalem, Israel; [†]Schneider Children's Hospital, Petach Tikva, affiliated to the Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; [‡]Pediatric Gastroenterology Unit, Soroka University Medical Center and Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel; [§]Pediatric Gastroenterology Institute, Ruth Children's Hospital, Rambam Medical Center, Haifa, Israel; [¶]Pediatric Gastroenterology Unit, Edmond and Lily Safra Children's Hospital, Sheba Medical Center and Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; [‖]Digestive diseases institute, Shaare Zedek Medical Center, Jerusalem, Israel; **Division of Gastroenterology, Rabin Medical Center, Petah Tikva Israel, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel; [††]Health Division, Maccabi Healthcare Services, Tel Aviv, Israel

Address correspondence to: Dan Turner MD, PhD, P.O.B 3235, Shmuel Bait, 12, Jerusalem 91031, Israel (turnerd@szmc.org.il).

doi: 10.1093/crocol/otaa027
Published online 24 April 2020

Currently, the Red Flag index, developed as part of an International Organization of IBD (IOIBD) initiative, is the major available tool to guide primary practitioners in selecting patients who are at risk of having IBD.[8] During the development of the Red Flag index, 21 items have been generated, of which 8 were retained by multivariate regression analysis. Items were weighted statistically according to the beta scores of the model (Table 2). Nonetheless, the Red Flag index has never been subjected to external validation and is not suitable neither for children nor for ulcerative colitis (UC).

We aimed to develop and validate a screening tool, named the IBD-REFER criteria, to guide general practitioners and pediatricians in selecting children and adults for early referral to a gastroenterologist for suspected CD and UC. We also aimed to compare the newly developed criteria with the existing Red Flags index.

METHODS

The IBD-REFER criteria have been developed by a combined clinimetric and psychometric approach, incorporating both clinical judgment by a Delphi group of experts and statistical analyses on retrospective cohorts of patients and controls. The criteria were validated on separate cohorts of patients and controls.

Stage 1: Delphi Group

Item generation, reduction, and formatting were initially performed judgmentally by a Delphi group of 10 IBD experts (5 pediatric and 5 adults) in an iterative process supplemented by a review of the literature. The group conducted a systemic literature review of signs and symptoms associated with early presentation of CD in the primary care setting. Thereafter, the experts added items based on clinical experience. Item reduction followed rank-ordering by the Delphi group in an iterative process until consensus has been reached.

Stage 2: Data-Driven Formatting and Weighting

The items obtained by the literature review and Delphi group were explored on 2 retrospective cohorts of adults and children (<18 years) referred to the out-patient clinics of Shaare Zedek Medical Center, Jerusalem, for gastrointestinal symptoms, during a 5-year period (2013–2018). The first cohort was composed of those eventually diagnosed with IBD (both CD and UC), and the other, of those whose final diagnosis was other than IBD (eg, IBS, food intolerance), serving as controls. The eligibility criteria were deliberately left wide to ensure a full breadth of possible IBD symptoms and signs, but patients who were referred with a specific concern clearly unrelated to IBD were excluded (eg, cancer surveillance) to mimic real-life referral challenges.

For each patient, we determined which of the generated signs and symptoms were reported at the time of referral. Further retrieved data included demography, family history, symptoms, laboratory results, as well as fecal calprotectin, anti-*Saccharomyces cerevisiae* antibodies (ASCA), and anti-neutrophil cytoplasmic antibodies (ANCA). We scored in all subjects the IOIBD Red Flag for allowing direct comparison.

Based on these 2 cohorts, we revised the IBD-REFER items both judgmentally (based on the patient data) and statistically (See Statistical paragraph) in an attempt to maximize sensitivity and specificity for differentiating the 2 cohorts. We removed uninformative items to maximize feasibility and reduce response burden, and, finally, we determined the optimal discriminative cutoff for each item (eg, optimal duration of the abdominal pain and optimal percent of weight loss) based on the data. The final draft was then sent to the Delphi group for comments and linguistic revisions.

Stage 3: External Validation

The final IBD-REFER, developed during the previous 2 steps, was then externally validated on 2 further retrospective cohorts of IBD cases and controls, enrolled with identical definitions and eligibility criteria as in the derivation stage. To ensure that the validation is indeed external, we enrolled the cohorts at other medical centers: the Rabin (adults) and Schneider's (children) Medical Centers in Petah Tikva, Israel.

Statistics

Variables are presented as means ± standard deviation or medians (interquartile range [IQR]) and compared by Student *t* test, Wilcoxon rank sum test, and $\times^2$ test as appropriate. Statistical selection of items utilized a regression model to determine the β-coefficient of each item to predict the correct diagnosis of IBD. Informative items were selected by Lasso procedure and stepwise logistic regression. We compared the sensitivity and specificity between the IBD-REFER and Red Flag criteria by McNemar test and area under the receiver operating characteristic (ROC) curve. For constructing the ROC curve, we calculated the number of positive items for each patient, when major criteria received 2 points and minor criteria 1 point. Analyses were performed using SPSS V15; *P*-values less than 0.05 were considered significant. This study was approved by the local ethical committee of all participating medical centers.

RESULTS

In the item-generation stage, 19 signs and symptoms were initially generated from the literature and the Delphi group process (Supplementary Table 1). The list was then reduced by rank-ordering to tentative 16 items, separated into 2 groups according to their perceived importance for early referral: 5 major items and 11 minor items that may require more than one to prompt a referral for suspected IBD (Table 1).

The final list determined by the Delphi group was then formatted mathematically in the data-driven stage. The medical charts of 200 IBD patients (100 adults and 100 children) and 165 controls were reviewed. As per the eligibility criteria, 65 patients who were referred for specific obvious reasons were excluded from the control group to ensure overlap with the IBD group (n = 25 celiac disease, n = 12 hepatitis, n = 5 lactose intolerance, and n = 23 peptic ulcer disease and others). Eventually, the non-IBD control group included 100 patients (50 children and 50 adults). Among the IBD group, 73 (73%) children and 78 (78%) adults were eventually diagnosed with CD, while 27 (27%) and 22 (22%) with UC. The mean age of children was 13.1 ± 3.5 years and of adults 30.4 ± 11.1 years. Forty-nine (49%) children and 49 (49%) adults were females. In the control group, the mean age was 9.5 ± 4.4 years in children and 38.2 ± 16 years in adults; the female gender rate was 50% and 54%, respectively.

TABLE 1. List of Tentative Items Determined by a Literature Search and the Delphi Group

Group 1 (Most Important—Any Item Is Sufficient for Early Referral)	Group 2 (Important—More Than One Item Is Required for Early Referral)
Diarrhea lasting for at least 1 month	Abdominal pain for at least 1 month
Bloody stool lasting for at least 1 week	Elevated serum inflammatory marker
Recurrent perianal abscess, fistula, or large inflamed skin tag	Weight loss (>10%)
Elevated fecal calprotectin	Iron deficiency anemia or hypoalbuminemia
Impaired growth in children	Fever of unknown origin
	Chronic fatigue
	First-degree family history
	Arthritis, uveitis, erythema nodosum, not meeting a clear rheumatological diagnosis
	Recurrent oral aphthous ulcerations
	Large chronic anal fissure
	Positive ASCA or ANCA

The draft of the judgmental criteria developed by the Delphi group (Table 1) performed well in this cohort (sensitivity/specificity: 100%/92% in children and 98%/76% in adults), better than the contending Red Flags index (Fig. 1). The median calprotectin level of the IBD patients was 370 µg/g (IQR 300–709). Nonetheless, in an attempt to improve further specificity, the criteria were revised based on the patient data. The following items were removed because they had β-coefficients less than 0.1 and seemed uninformative in manual simulations: hypoalbuminemia, iron deficiency anemia, fever of unknown origin, fatigue, rheumatologic symptoms without a rheumatological diagnosis, recurrent oral aphthous ulcerations, large chronic fissures, and large inflamed skin tags (Table 2). Removal of these items does not mean that they are not important in diagnosing IBD but that, in practice, these are typically accompanied by more common signs and symptoms and thus redundant. We also moved calprotectin and growth delay from the major to the minor criteria group because it seemed to improve the specificity without altering the sensitivity.

The final IBD-REFER criteria were composed of 3 major criteria and 7 minor criteria (Fig. 2) and had very high sensitivity and specificity on the derivation cohort (98%/94% in children and 94%/88% in adults).

We next validated the criteria on the external cohorts, which included 100 IBD patients (50 children and 50 adults) and 85 non-IBD patients. From the comparison group we excluded 35 patients referred for specific diagnoses (ie, celiac disease, n = 5; liver dysfunction, n = 7; lactose intolerance, n = 2; and peptic ulcer disease and others, n = 21). Eventually, the non-IBD control group included 50 patients (25 children and

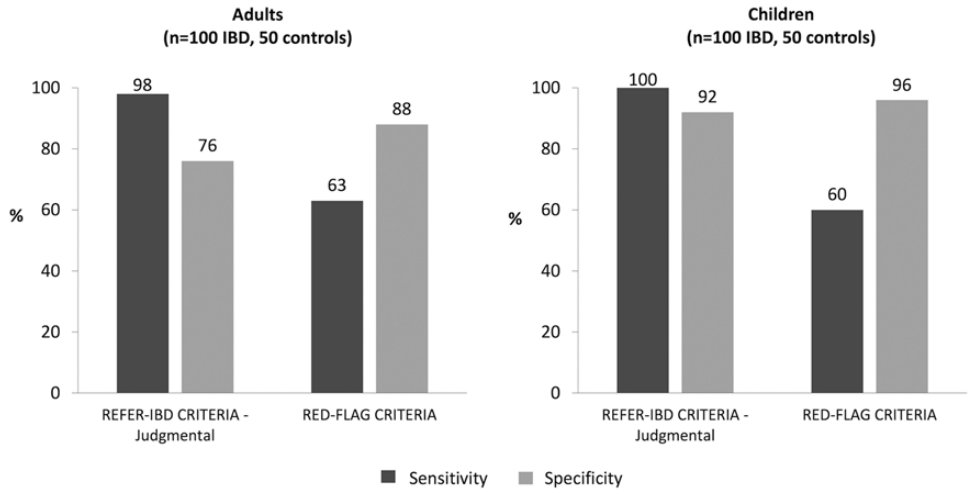

FIGURE 1. Sensitivity and specificity of the "judgmental version" IBD-REFER criteria and the Red Flag index of an International Organization of IBD (IOIBD) in the development cohort to differentiate IBD from non-IBD patients at referral.

25 adults). Among the IBD group, 37 (74%) children and 24 (48%) adults had CD, while 13 (26%) and 26 (52%) had UC. The mean age of the pediatric and adult cohorts was 13.7 ± 3.1 and 43.3 ± 22.4 years, respectively; 46% and 64% were females.

TABLE 2. Items Coefficient by Multiple Regression

Item	Coefficient
Bloody stool	0.48
perianal abscess/fistula	0.31
Diarrhea	0.28
Elevated serum inflammatory marker	0.26
First-degree family history	0.26
Elevated fecal calprotectin	0.21
Rheumatological symptoms	0.19
Impaired growth in children	0.14
Weight loss	0.13
Iron deficiency anemia or hypoalbuminemia	0.11
Chronic fatigue	0.09
Large chronic anal fissure	0.09
Positive ASCA or ANCA	0.07
Chronic abdominal pain	0.02
Fever of unknown origin	−0.03
Recurrent oral aphthous ulcerations	−0.06

In the control group, the corresponding mean ages were 12.1 ± 4.2 and 61.1 ± 17.5; 40% and 36% were females.

As in the development cohort, the final IBD-REFER criteria had very high sensitivity and specificity in both children and adults (96%/96% in children and 98%/96% in adults, respectively), significantly higher than achieved by the IOIBD Red Flag index (Fig. 3). Similarly, the area under the ROC curve was significantly higher in IBD-REFER criteria compared with the IOIBD Red Flag criteria (AUC = 0.97 [95% CI 0.92–0.99] vs 0.78 [95%CI 0.7–0.86]; $P < 0.001$) (Fig. 4).

Finally, in a sensitivity analysis, we explored the utility of 2 simplified versions of the IBD-REFER criteria, one without the laboratory items (ie, fecal calprotectin, ESR/CRP, ASCA/ANCA) and the other with the major criteria only (ie, excluding all items on the minor list). The performance of both simplified versions was inferior to the full version: the sensitivity/specificity of the former was 89%/97% in children and 91%/93% in adults and of the latter 76%/99% in children and 87%/95% in adults.

DISCUSSION

We have developed and validated a simple screening tool based on routinely recorded signs and symptoms to guide primary practitioners in selecting children and adults for early referral to the gastroenterologist for suspected CD and UC. Our

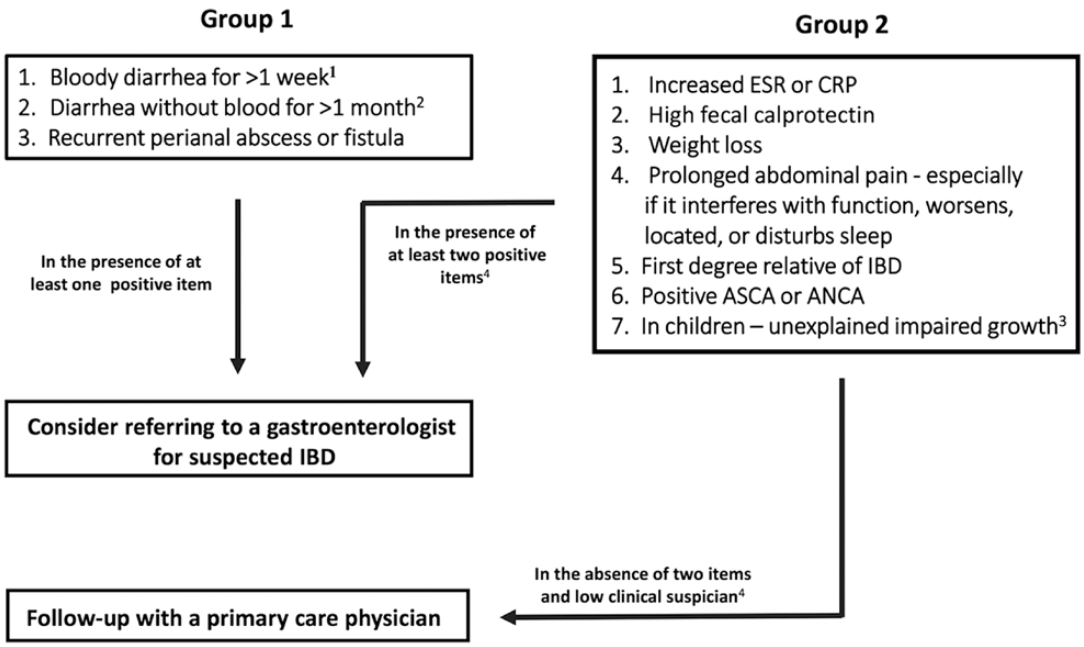

FIGURE 2. The final IBD-REFER criteria. 1. Bloody stools without diarrhea not associated with a fissure or hemorrhoids also require early referral to a gastroenterologist, but in this case, not necessarily for suspected IBD. 2. Particularly in the event of nocturnal diarrhea or rectal urgency. 3. Normal serology to celiac disease should be confirmed. If height is more impaired than weight in the absence of gastrointestinal symptoms, a referral to an endocrinologist should be considered. 4. When clinical suspicion is high, the presence of even one Group 2 item is sufficient for a referral to the gastroenterologist or for determining fecal calprotectin.

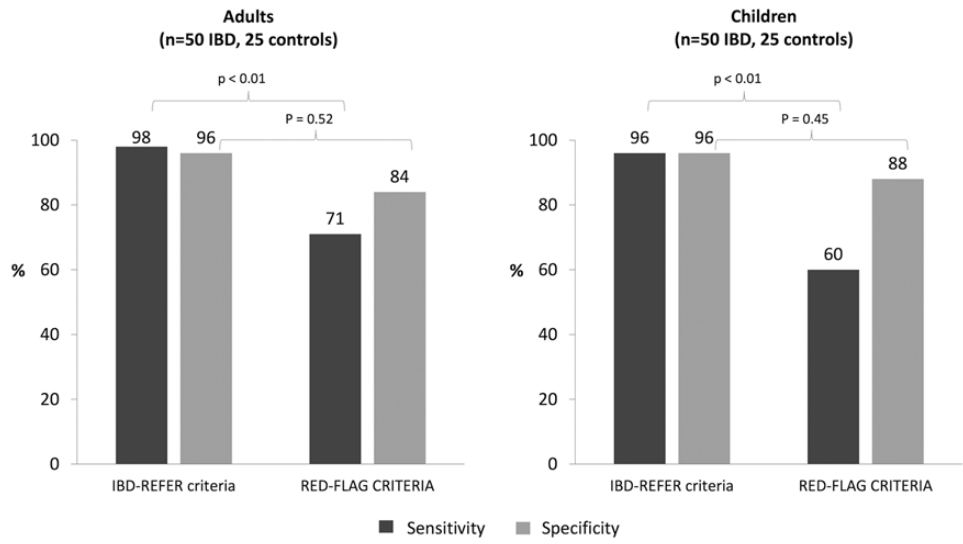

FIGURE 3. Sensitivity and specificity of the final IBD-REFER criteria compared with the Red Flags index, in the external validation cohort.

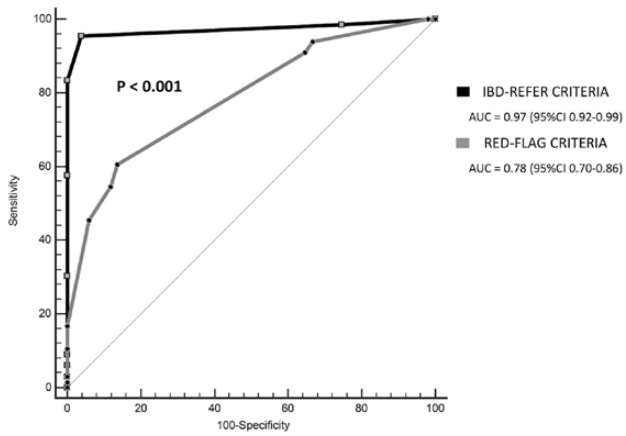

FIGURE 4. Receiver operating characteristic curve of IBD-REFER criteria and Red Flag index.

overall goal was to decrease the time elapsed from the onset of symptoms to the diagnosis of IBD, thereby allowing early appropriate treatment and reducing the risk of disease-related complications. The IBD-REFER criteria seemed to perform better than the Red Flag index and, unlike the latter, is relevant not only for CD and adults but also for UC and children.[9]

We explored 2 simplified versions but these were not endorsed given the attempt to maximize both sensitivity and specificity. In developing referring criteria, ensuring high specificity is of utmost importance to avoid over-utilization of scarce resources. Low specificity will inevitably lead to the referral of many patients with gastrointestinal symptoms, but who are at low risk of having IBD. In turn, this will lower the confidence of primary physicians in the criteria. Nonetheless, in our study, the high specificity was not at the expense of sensitivity which was also very high (96% in children and 98% in adults).

The incorporation of fecal calprotectin into the algorithm was not intuitive. On the one hand, it has become an important screening test for IBD; but on the other hand, it is not commonly used by primary care physicians and its low specificity may impact the accuracy of the criteria.[10] Thus, calprotectin is a supplementary rather than mandatory item, allowing its incorporation when performed. In addition, because some of the items in our final criteria are not specific to IBD (eg, abdominal pain, diarrhea, and impaired growth), other possible diagnoses should be considered by the primary care physician such as lactose intolerance, enteric infections, celiac disease, and growth hormone deficiency.

The median delay from symptoms to diagnosis range in western countries is from 4 to 9 months, and approximately 25% of patients are diagnosed more than 2 years after the onset of symptoms.[5] In Asia, the median delay may be as long as 18 months, and 37% of patients are diagnosed with a more than 2-year delay.[6] In turn, the diagnostic delay has been consistently associated with an increased risk of disease-related complications.[11–13] Several studies have shown that diagnostic delay is associated with an increased risk of bowel damage, intestinal resection,[11, 12, 14] and, in children, growth impairment.[13]

The strengths of our study include the use of robust combined judgmental-clinimetric and mathematical strategies, the use of 2 separate cohorts enrolled at different hospitals but with identical eligibility criteria, and the inclusion of both children and adults because many general practitioners treat both. We also used a challenging comparison group composed of patients with symptoms requiring gastroenterology consultation but without those with an obvious working diagnosis. Nonetheless, the study is limited by its retrospective design and, consequently, the inherent inability to report predictive values. We enrolled patients already referred to a gastroenterologist

and not from primary care clinics. However, all IBD patients eventually see a gastroenterology and by enrolling consecutive patients we practically included also those who were referred with significant delays capturing the full breadth of patients. This, because even if a patient is missed in the screening by the primary care physician he/she will eventually be seen in the clinic.

The newly developed IBD-REFER criteria set is an accurate tool for selecting patients to expedite gastrointestinal consultation for suspected IBD. A Swiss study showed that time from the gastroenterologist visit to the diagnosis of CD was also a contributing factor in the diagnostic delay and not only from symptoms to referral.[11] To that end, the IBD-REFER criteria can be used also by the gastroenterology units for prioritizing cases for more urgent visits and endoscopic assessment. In the era of "early treatment" recommendation for high-risk patients, shortening the time from onset of symptoms to diagnosis may impact long-term important outcomes of IBD patients.

SUPPLEMENTARY MATERIAL

Supplementary data are available at *Crohn's & Colitis 360* online.

Conflict of interest statement. O.A., A.S., R.L.T., B.Y., R.S., B.K., R.K., and E.L. declared no conflicts of interest. G.F. received the consultation fee from Eli Lilly and Abbvie. O.L. received a travel grant from Janssen. A.A. obtained research grants from Abbvie and Janssen; consultation and lectures fee from Abbvie. D.S. received in the last 3 years consulting and lecturing fees from Abbvie and research grant from Takeda. A.B. received consultation fee, advisory board, speakers bureau, research grant, or honorarium from Takeda, Janssen, AbbVie, Neopharm, and Pfizer. I.D. received the last 3 years consultation fee, research grant, or honorarium from Janssen, Pfizer, Abbvie, Takeda, Genentech/Roche, Neopharm, Ferring, Rafa Laboratories, Falk Pharma, Nestle, Given Imaging/Medtronic, Gilead, Celgene, Arena, Sublimity, Celltrion, and Altman Research. D.T. received the last 3 years consultation fee, research grant, royalties, or honorarium from Janssen, Pfizer, Hospital for Sick Children, Ferring, Abbvie, Takeda, Biogen, Neopharm, Uniliver, Atlantic Health, Shire, Celgene, Lilly, and Roche.

REFERENCES

1. Lakatos PL. Recent trends in the epidemiology of inflammatory bowel diseases: up or down? *World J Gastroenterol.* 2006;12:6102–6108.
2. 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.
3. Danese S, Semeraro S, Papa A, et al. Extraintestinal manifestations in inflammatory bowel disease. *World J Gastroenterol.* 2005;11:7227–7236.
4. Colia R, Corrado A, Cantatore FP. Rheumatologic and extraintestinal manifestations of inflammatory bowel diseases. *Ann Med.* 2016;48:577–585.
5. Vavricka SR, Spigaglia SM, Rogler G, et al.; Swiss IBD Cohort Study Group. Systematic evaluation of risk factors for diagnostic delay in inflammatory bowel disease. *Inflamm Bowel Dis.* 2012;18:496–505.
6. Banerjee R, Pal P, Girish BG, Reddy DN. Risk factors for diagnostic delay in Crohn's disease and their impact on long-term complications: how do they differ in a tuberculosis endemic region? *Aliment Pharmacol Ther.* 2018;47:1367–1374.
7. Walters TD, Kim MO, Denson LA, et al.; PRO-KIIDS Research Group. Increased effectiveness of early therapy with anti-tumor necrosis factor-α vs an immunomodulator in children with Crohn's disease. *Gastroenterology.* 2014;146:383–391.
8. Danese S, Fiorino G, Mary JY, et al. Development of red flags index for early referral of adults with symptoms and signs suggestive of Crohn's disease: an IOIBD initiative. *J Crohns Colitis.* 2015;9:601–606.
9. Chmiel C, Vavricka SR, Hasler S, et al. Feasibility of an 8-item questionnaire for early diagnosis of inflammatory bowel disease in primary care. *J Eval Clin Pract.* 2019;25:155–162.
10. Chmiel C, Senn O, Hasler S, et al. Can the calproquest predict a positive calprotectin test? A prospective diagnostic study. *PLoS One.* 2019;14:e0224961.
11. Schoepfer AM, Dehlavi MA, Fournier N, et al.; IBD Cohort Study Group. Diagnostic delay in Crohn's disease is associated with a complicated disease course and increased operation rate. *Am J Gastroenterol.* 2013;108:1744–53; quiz 1754.
12. Zaharie R, Tantau A, Zaharie F, et al.; IBDPROSPECT Study Group. Diagnostic delay in Romanian patients with inflammatory bowel disease: risk factors and impact on the disease course and need for surgery. *J Crohns Colitis.* 2016;10:306–314.
13. Ricciuto A, Fish JR, Tomalty DE, et al. Diagnostic delay in Canadian children with inflammatory bowel disease is more common in Crohn's disease and associated with decreased height. *Arch Dis Child.* 2018;103:319–326.
14. Nahon S, Lahmek P, Paupard T, et al. Diagnostic delay is associated with a greater risk of early surgery in a French cohort of Crohn's disease patients. *Dig Dis Sci.* 2016;61:3278–3284.