**TITLE:** Recommendations for Standardizing Magnetic Resonance Imaging-based Evaluation of Perianal Fistulizing Disease Activity in Pediatric Crohn's Disease Clinical Trials

**AUTHORS:** Eileen Crowley, MD,[1,2] Christopher Ma, MD MPH,[2,3,4] Leonard Guizzetti, PhD,[2] Guangyong Zou, PhD, [2,5] Peter J. Lewindon, MD,[6,7] Michael S. Gee, MD, PhD,[8,9] Jeffrey S. Hyams, MD,[10] Michael J. Rosen, MD, MSCI[11] Daniel von Allmen, MD,[12,13] Anthony deBuck van Overstraeten, MD, MsC[14] Lisa M. Shackelton, PhD,[2] Julie Remillard, MSc, [2] Lauren Schleicher, [2] Jonathan R. Dillman, MD, MSc[15,16] Jordi Rimola, MD, PhD,[17] Stuart A. Taylor, MD,[18] Joel G. Fletcher, MD,[19] Peter C. Church, MD,[20] Brian G. Feagan, MD,[2,5,22] Anne M. Griffiths, MD,[21] Vipul Jairath, MD,[2,5,22] Mary-Louise C. Greer, MBBS[23]

**AUTHOR AFFILIATIONS:**
[1]Department of Pediatrics, Division of Pediatric Gastroenterology, Children's Hospital Western Ontario, Western University, London Health Sciences Centre, London, Ontario, Canada
[2]Alimentiv Inc., London, ON, Canada
[3]Division of Gastroenterology & Hepatology, Department of Medicine, University of Calgary, Calgary, AB, Canada
[4]Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[5]Department of Epidemiology and Biostatistics, Western University, London, Ontario, Canada
[6]Department of Gastroenterology, Queensland Children's Hospital, University of Queensland, Brisbane, Queensland, Australia
[7]Children's Medical Research Institute, University of Queensland, Brisbane, Queensland, Australia
[8]Department of Radiology, Massachusetts General Hospital Boston, MA, USA
[9]Department of Radiology, Harvard Medical School, Boston, MA, USA
[10]Division of Digestive Diseases, Hepatology, and Nutrition, Connecticut Children's Medical Center, Hartford, CT, USA
[11]Division of Pediatric Gastroenterology, Hepatology and Nutrition, Department of Pediatrics, Stanford Medicine, Stanford, CA, USA
[12]Department of Pediatric Surgery and Surgical Services, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA
[13]Surgical Services, Cincinnati Children's Hospital Medical Center, Department of Surgery, University of Cincinnati, Cincinnati, OH, USA
[14]Department of Surgery, Mount Sinai Hospital, University of Toronto, Toronto, ON, Canada
[15]Department of Radiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA
[16]Department of Radiology, University of Cincinnati College of Medicine, Cincinnati, OH, USA
[17]IBD Unit, Radiology Department, Hospital Clinic Barcelona, IDIBAPS, Spain
[18]Centre for Medical Imaging, Charles Bell House, University College London, London, UK
[19] Department of Radiology, Mayo Clinic, Rochester, MN, USA
[20]Division of Gastroenterology, Hepatology & Nutrition, Department of Paediatrics and IBD Centre, SickKids Hospital, Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada
[21]Division of Gastroenterology, Hepatology & Nutrition, Department of Paediatrics and IBD Centre, SickKids Hospital, University of Toronto, Toronto, ON, Canada

[22]Department of Medicine, Division of Gastroenterology, Western University, London, ON, Canada
[23]Department of Diagnostic Imaging, the Hospital for Sick Children, Department of Medical Imaging, University of Toronto, Toronto, ON, Canada

**CORRESPONDENCE:** Dr Eileen Crowley; Division of Pediatric Gastroenterology, Hepatology & Nutrition, Children's Hospital Western Ontario, London Health Sciences Center; Assistant Professor, Western University, 800 Commissioners Rd E, London, ON N6A 5W9; Tel: 519-685-8177; Fax: 519-685-8105; Email: eileen.crowley@lhsc.on.ca

**FUNDING:** None

**FINANCIAL DISCLOSURES & CONFLICTS OF INTEREST:**

**EC** has received an educational grant and speaker fees from AbbVie, Pfizer, and consulting fees from Alimentiv Inc.

**CM** has received consulting fees from AbbVie, Alimentiv, Amgen, AVIR Pharma Inc, BioJAMP, Bristol Myers Squibb, Celltrion, Ferring, Fresenius Kabi, Janssen, McKesson, Mylan, Takeda, Pendopharm, Pfizer, Roche; speaker's fees from AbbVie, Amgen, AVIR Pharma Inc, Alimentiv, Bristol Myers Squibb, Ferring, Fresenius Kabi, Janssen, Takeda, Pendopharm, and Pfizer; royalties from Springer Publishing; research support from Ferring, Pfizer

**LG** is an employee of Alimentiv Inc.

**GYZ** has received consulting fees from Alimentiv Inc.

**PJL** has received speakers fees from Janssen, Abbvie and Fering.

**MG** has research agreements with Siemens Healthineers, General Electric Healthcare, and Motilent, LLC.

**JSH** is an advisory board member for Janssen, Pfizer, Lilly, Bristol Myers Squibb and a consultant for Takeda and Thetis.

**MJR** has served on an advisory board for Pfizer.

**DvA** has no relevant disclosures.

**AdBvO** has received consulting fees from Takeda.

**LMS** has received consulting fees from Alimentiv Inc.

**JR** is an employee of Alimentiv Inc.

**LS** is an employee of Alimentiv Inc.

**JRD** has grant/research support from Motilent, Bracco Diagnostics, Philips Healthcare, GE Healthcare, and Perspectum Ltd.

**JR** has received grants from Abbvie and consulting fees from Boehringer Ingelheim, Janssen, Origo, Takeda and Alimentiv Inc.

**SAT** has received consulting fees from Alimentiv Inc and hold shares in Motilent, LLC.

**JGF** has received grant support with funds to his institution from Siemens Healthineers, Leona M. and Harry B. Helmsley Charitable Trust, Alimentiv Inc., and consulting with funds to institution from Genentech, Boehringer Ingelheim, Glaxo Smith Kline, Janssen, Medtronic, and Takeda.

**PC**

**BGF** has received grant/research support from AbbVie Inc., Amgen Inc., AstraZeneca/MedImmune Ltd., Atlantic Pharmaceuticals Ltd., Boehringer-Ingelheim, Celgene Corporation, Celltech, Genentech Inc/Hoffmann-La Roche Ltd., Gilead Sciences Inc., GlaxoSmithKline (GSK), Janssen Research & Development LLC., Pfizer Inc., Receptos Inc./Celgene International, Sanofi, Santarus Inc., Takeda Development Center Americas Inc., Tillotts Pharma AG, and UCB; consulting fees from Abbott/AbbVie, Akebia Therapeutics, Allergan, Amgen, Applied Molecular Transport Inc., Aptevo Therapeutics, Astra Zeneca, Atlantic Pharma, Avir Pharma, Biogen Idec, BioMx Israel, Boehringer-Ingelheim, Bristol-Myers Squibb, Calypso Biotech, Celgene, Elan/Biogen, EnGene, Ferring Pharma, Roche/Genentech, Galapagos, GiCare Pharma, Gilead, Gossamer Pharma, GSK, Inception IBD Inc, JnJ/Janssen, Kyowa Kakko Kirin Co Ltd., Lexicon, Lilly, Lycera BioTech, Merck, Mesoblast Pharma, Millennium, Nestle, Nextbiotix, Novonordisk, Pfizer, Prometheus Therapeutics and Diagnostics, Progenity, Protagonist, Receptos, Salix Pharma, Shire, Sienna Biologics, Sigmoid Pharma, Sterna Biologicals, Synergy Pharma Inc., Takeda, Teva Pharma, TiGenix, Tillotts, UCB Pharma, Vertex Pharma, Vivelix Pharma, VHsquared Ltd., and Zyngenia; speakers bureau fees from Abbott/AbbVie, JnJ/Janssen, Lilly, Takeda, Tillotts, and UCB Pharma; is a scientific advisory board member for Abbott/AbbVie, Allergan, Amgen, Astra Zeneca, Atlantic Pharma, Avaxia Biologics Inc., Boehringer-Ingelheim, Bristol-Myers Squibb, Celgene, Centocor Inc., Elan/Biogen, Galapagos, Genentech/Roche, JnJ/Janssen, Merck, Nestle, Novartis, Novonordisk, Pfizer, Prometheus Laboratories, Protagonist, Salix Pharma, Sterna Biologicals, Takeda, Teva, TiGenix, Tillotts Pharma AG, and UCB Pharma; and is the Senior Scientific Officer of Alimentiv Inc.

**AMG** has received consulting fees from Abbvie, Amgen, Bristol Myers Squibb, Lilly, Janssen, Merck, Mylan, Pfizer; speaker fees from Abbvie, Janssen, Nestle; investigator-initiated research support from Abbvie.

**MLG** has received speaker fees from Abbvie and Samsung; investigator-initiated research support from Abbvie.

**VJ** has received **consulting/advisory board fees** from AbbVie, Alimentiv Inc, Arena pharmaceuticals, Asahi Kasei Pharma, Asieris, Astra Zeneca, Bristol Myers Squibb, Celltrion, Eli Lilly, Ferring, Flagship Pioneering, Fresenius Kabi, Galapagos, GlaxoSmithKline, Genentech, Gilead, Janssen, Merck, Metacrine, Mylan, Pandion, Pendopharm, Pfizer, Protagonist, Prometheus, Reistone Biopharma, Roche, Sandoz, Second Genome, Sorriso pharmaceuticals, Takeda, Teva, Topivert, Ventyx, Vividion; **speaker's fees** from, Abbvie, Ferring, Bristol Myers Squibb, Galapagos, Janssen Pfizer Shire, Takeda, Fresenius Kabi

Alimentiv Inc. is an academic gastrointestinal contract research organization, operating under the Alimentiv Health Trust. Alimentiv Inc. provides comprehensive clinical trial services, precision medicine offerings, and centralized imaging solutions for endoscopy, histopathology, and other imaging modalities. The beneficiaries of the Alimentiv Health Trust are the employees of the enterprises it holds. EC, CM, GYZ, LMS, BGF, and VJ are consultants to Alimentiv Inc.; they do

not hold equity positions or shares in Alimentiv Inc. EC, CM, GYZ, BGF, and VJ have a primary academic appointment.

**SUMMARY:**

This study determined the appropriateness of MRI-based disease activity indices, methods for assessment, and clinical trial design for pediatric perianal fistulizing CD. These recommendations will aid index development for assessing disease activity in future clinical studies.


**ABSTRACT**

**Background:** Perianal fistulas and abscesses are common complications in children with Crohn's disease (CD). A validated imaging assessment tool for quantification of perianal disease severity and activity is needed to evaluate treatment response. We aimed to identify magnetic resonance imaging (MRI)-based measures of perianal fistulizing disease activity and study design features appropriate for pediatric patients.

**Methods**: Seventy-nine statements relevant to MRI-based assessment of pediatric perianal fistulizing CD activity and clinical trial design were generated from literature review and expert opinion. Statement appropriateness was rated by a panel (N=15) of gastroenterologists, radiologists and surgeons using modified RAND/University of California Los Angeles appropriateness methodology.

**Results**: The modified Van Assche Index (mVAI) and the Magnetic Resonance Novel Index for Fistula Imaging in CD (MAGNIFI-CD) were agreed to be appropriate instruments for use in pediatric perianal fistulizing disease clinical trials. Although there was concern regarding the use of intravascular contrast material in pediatric patients, its use in clinical trials was considered appropriate. Appropriate trial inclusion criteria included a clinically evident fistula tract and

radiologic disease defined as at least one fistula or abscess on pelvic MRI. A co-primary clinical and radiologic endpoint and inclusion of a patient-reported outcome were also considered appropriate.

**Conclusion**: Outcomes of treatment of perianal fistulizing disease in children must include MRI, and existing multi-item measures, specifically the mVAI and MAGNIFI-DC can be adapted and used for children. Reliability and validity assessment specifically in pediatric patients is planned.

**KEYWORDS:** pediatric; Crohn's disease; perianal fistula; clinical trial; pelvic MRI

**KEY MESSAGES:**

**What is already known?**

A validated imaging assessment tool for quantification of disease activity and severity may aid evaluation of treatment response for new interventions for perianal fistulas in children with Crohn's disease.

**What is new here?**

We identified appropriate magnetic resonance imaging-based measures of disease activity and study design features.

**How can this study help patient care?**

These recommendations will aid index development, standardization of trial outcomes, and investigation of new treatments.

## INTRODUCTION

Perianal fistulizing disease is a disabling complication of pediatric Crohn's disease (CD) and is characterized by pain, purulent discharge, incontinence, and impaired quality of life. In a Canadian multicenter prospective study of 1,092 patients with new onset inflammatory bowel disease (IBD), of the 63% patients with CD, 16% had perianal fistulizing disease as defined by the Paris classification of pediatric IBD.[1] The fundamental goals for treatment of perianal fistulizing CD in children are control of local sepsis and induction of mucosal and fistula tract healing, which frequently requires a combined medical-surgical approach. Consensus guidelines by all academic societies recommend tumor necrosis factor (TNF) antagonists as the initial choice of treatment for patients with signicant perianal fistulizing disease.[2-3] Surgical interventions may be considered, with small uncontrolled studies showing benefit of seton placement and fistulotomy in selected patients,[4] however efficiacy of surgical intervention must be weighed against the risk of fecal incontinence inherent to surgery involving the sphincters. A systematic review including four studies that examined the efficacy of infliximab in pediatric patients reported clinical resolution of perianal disease in 110 of 197 patients (55%), and a partial response to treatment in a further 33 (17%).[4] Despite clinical cessation of drainage, however, subclinical disease may persist that is undetectable by visual or physical examination.[5] Thus, other diagnostic modalities are needed to accurately assess perianal fistulizing disease activity.

Pelvic magnetic resonance imaging (MRI) targeted to the perianal region is currently considered the best non-invasive tool for assessing perianal fistulas, and has been shown to have greater sensitivity for detecting perianal disease than clinical evaluation.[6] The Parks[7] and St. James' University Hospital[8] classifications are historical surgical- and MRI-derived anatomical methods for classifying perianal disease (Supplemental Table 1), and are used clinically to describe the

anatomic relationship of the fistula tract(s) to the anal sphincter complex but do not provide an estimation of the severity of activity nor prognostic value for associated outcomes. Of MRI-based measures aiming to quantify perianal inflammation, the Van Assche Index (VAI)[9] has the most reported use in the literature to date but its responsiveness is suboptimal.[10] A modified version of this index was developed in 2017 (Supplemental Table 2) based on the reliability and responsiveness of disease descriptors.[10] The Magnetic Resonance Novel Index for Fistula Imaging in CD (MAGNIFI-CD) was more recently developed and internally validated for evaluating perianal fistulizing disease activity in adults with CD (Supplemental Table 3).[11] MAGNIFI-CD assesses 6 weighted MRI items (number of fistula tracts, hyperintensity of primary tract on post-contrast T1-weighted images, dominant feature, fistula length, extension, and inflammatory mass) to calculate an overall score of fistulizing CD activity. This index was shown to be reliable and responsive to change.[11] The performance of MAGNIFI-CD in the pediatric population, where the morphology of perianal disease may be different, is currently unknown. The Pediatric MRI-based Perianal Crohn's disease (PEMPAC) index for pediatric CD has also recently been developed (Supplemental Table 4) and validated. As with MAGNIFI-CD in adults, PEMPAC was shown to be reliable and responsive to change.[12]

There is a need for safe and effective therapies for the management of fistulizing CD in the pediatric population as well as in adults. A multi-item measure is required to assess response, and may also offer potential for prognostication.[13] It would be advantageous to use the same instrument across all age groups, even if some modification is required for children. We therefore employed Research and Development/University of California Los Angeles (RAND/UCLA) appropriateness methodology[14] to determine the appropriateness of existing MRI-based fistulizing CD activity indices for assessment of pediatric perianal fistulizing CD; methods for assessment and/or

definitions of specific features of pediatric perianal fistulizing disease activity; and aspects of clinical study design for pediatric fistulizing CD. We herein report the first stage in ongoing studies to assess the reliability and validity of MR-based measures of perianal disease activity in children.

## MATERIALS AND METHODS

### Statement generation

A list of statements for determination of appropriateness was drafted based on literature review and expert opinion that included MRI indices (and index component items) currently employed for assessment of CD activity, definitions and/or scoring methods for assessment of disease activity, and elements of study design relevant to clinical trials in pediatric perianal fistulizing CD (including trial inclusion criteria and outcome assessment) (EC, LMS, MLG).

### Recruitment of panelists

An international, multidisciplinary panel of 15 experts (7 pediatric and adult gastroenterologists, 2 surgeons, 6 pediatric and adult radiologists) were invited to participate. Panelists were selected based on publication record, clinical and/or research expertise in pediatric gastroenterology, radiology, surgery and/or experience in index development or methodology. These criteria took precedence over geographical representation. Panel participant selection was performed by MLG and VJ after review of a list of potential experts meeting the criteria described above.

A modified RAND/UCLA appropriateness methodology was used to assess the face validity (i.e., the extent to which an item appears to address the concept it purports to measure) and feasibility of statements.[15] This methodology, which is widely accepted, iterative, and evidence-based,

employs a modified Delphi panel approach to combine the best available evidence with the clinical experience of relevant experts, without forcing consensus.[16]

**First panel meeting/survey and analysis of survey results**

The initial list of statements was circulated to panelists for review prior to the introductory meeting, which was conducted virtually in June 2020. Assumptions for the development of the list and the individual statements were discussed with the panel during the moderated teleconference, and the list was modified based on panel feedback. Statements on the modified list were then incorporated into a survey that was circulated to the panelists for online voting.

The appropriateness of each statement was rated anonymously by each panelist on a scale from 1 to 9 (1 = highly inappropriate, 9 = highly appropriate). Panelists were also permitted to add free text comments. Statement appropriateness was assessed as described in the RAND manual and based on the median panel rating and degree of panel disagreement, where statements with median scores from 1 to ≤3.5 without disagreement were considered inappropriate, those with median scores >3.5 to <6.5 or any median score with disagreement were considered uncertain, and statements with median scores ≥6.5 to 9 without disagreement were appropriate.[17] Disagreement was defined as having at least 5 rating(s) in each of the lowest (1 to 3.5) and highest (6.5 to 9) median ranges, and was based on recommendations for panels with 15 experts. Disagreement is not explicitly defined for larger panels in the RAND/UCLA manual, but rather implicitly defines disagreement as having at least one-third of the panellists rating in each of the extreme median groups.[17]

**Second panel meeting/survey**

Results of the first survey were distributed to panelists and presented in a moderated teleconference. Areas of disagreement regarding item appropriateness were identified and panelists were asked to explain the rationale behind their responses. In accordance with RAND/UCLA appropriateness methodology, no attempt was made to force the panel to consensus. The survey was revised based upon the second panel meeting and a second survey was circulated for rating by the panelists. Statement appropriateness on the second survey was analysed as described above.

## RESULTS

### Statement generation and survey

The draft of the statement list was arranged according to the relevant topics (Figure 1). Table 1 summarizes the component items from existing disease activity indices that were considered. A final survey consisting of 79 statements was circulated. Of these, 59 were rated as appropriate, 1 rated as inappropriate and 19 were rated as uncertain (Table 2). Uncertain statements were further reviewed, and ratings were tabulated according to specialty (radiologist versus gastroenterologists and surgeons). No meaningful subjective differences were evident, and most statements remained uncertain regardless of the survey responder's specialty. Statements rated as uncertain by the panel whose appropriateness rating differed by specialty are shown in Supplemental Table 5.

### Appropriateness of disease activity indices

The MAGNIFI-CD and mVAI were considered appropriate for assessment of disease activity in pediatric perianal fistulizing CD clinical trials, however the panel was uncertain regarding the use of the original VAI for this purpose. The fact that the VAI does not include assessment of fistula

size or account for the presence of either anovaginal fistulas, horseshoe extensions, or an inflammatory mass was a concern for some panelists. Whilst the St. James' University Hospital and Parks Classifications were considered useful for the assessment of anatomy and surgical risk, they were considered inadequate for assessment of disease activity and do not differentiate outcome characteristics.

**Appropriateness of radiologic items**

Component items of existing indices were overall rated as appropriate for assessment of disease activity in pediatric perianal fistulizing CD clinical trials. Although additional items were proposed during the RAND process, the panel was uncertain whether assessment of aggregate fistula length, sinuses, number of internal openings and extra-anal regional inflammation was appropriate in clinical trials of perianal fistulizing CD.

The assessment of length of the predominant fistula was rated as appropriate by the panel. However, the panel raised potential concerns regarding the assessment of fistula length in clinical trials, including the accuracy of measurement on 2D images, as well as confounding due to differential healing rates. The panel additionally questioned the portion of the tract that should be measured on follow-up (i.e., measuring the entire tract including fibrotic components versus measuring only T2 hyperintense portions reflecting granulation tissue and/or fluid). Although assessment of number of fistula tracts was considered appropriate and a marker of disease severity and/or extent, panelists advised that this feature would be unlikely to change in response to currently available therapy and that number of fistulas may be a more useful radiological measure of disease characteristics rather than response to therapy. The panel advised that accurate assessment of number of fistula tracts may be limited by visualization of internal anatomy that may be less well defined in pediatric patients relative to adults, delineation influenced by the

homogeneity of fat suppression, and small size of the anal canal and fistula (or its ramifications) in this population. Panelists further noted that a disconnect often occurs between what is visualized on the perineum by the clinician and radiologic findings. Questions were also raised regarding whether this assessment should be limited to tracts with evidence of activity, or to those that reach the skin.

Uncertainty regarding assessment of the number of internal openings, which some panelists equated to evaluation of fistula, may have arisen from the difficulty of assessment, as internal openings are not always well seen on MRI, and can be affected by the presence of a seton. Importantly, items responsive to change with a seton in place were considered appropriate for radiologic assessment of pediatric fistulizing CD activity. The panel was uncertain regarding the need to differentiate sinus tracts from fistulas.

**Methods for scoring**

Although the component items of existing instruments were generally considered appropriate, there was panel uncertainty regarding the methods for scoring some of the features, including fistula extension, abscess volume and rectal wall involvement. There was considerable discussion about the assessment and scoring of fistula extensions, particularly regarding horseshoe extensions, bidirectional ramifications that are often but not always defined as crossing midline.[18, 19] Not all panelists agreed with the weighting of horseshoe and infralevatoric extensions or that the two were mutually exclusive. Although scoring of fistula extension in the MAGNIFI-CD index applies the same weight to scoring of infra- and supralevatoric extensions, some panelists believed these should be scored separately.

The panelists agreed that T1-weighted post-contrast images should be fat-suppressed. Although the scoring of items in the inflammatory mass category according to MAGNIFI-CD conventions

was considered appropriate, the panel noted potential difficulties in distinguishing focal lesions from collections, and some panelists believed that the mixing of lesion type in this item was inconsistent with the notion that inflammatory masses are a distinct entity from fluid containing abscesses. Finally, there were suggestions that assessment of this feature should be measured on more than one axis, rather than the shortest axis, as prescribed in MAGNIFI-CD.

Although assessment of rectal wall involvement was considered appropriate, scoring this item consistent with the original VAI (e.g., normal versus thickened) was rated as uncertain, as some panelists suggested that more precision was required, and they noted imprecision in absolute wall thickness, changing with the degree of rectal distension. Some panelists proposed assessment of proctitis (defined per mVAI as increased wall thickness and size of mesorectal lymph nodes (>5mm), fibrofatty proliferation, increased mural or perimural T2 signal and enhancement) in preference to rectal wall involvement alone (defined per VAI as normal or thickened).

**Clinical trial inclusion and outcome criteria**

Appropriate inclusion criteria for clinical trials evaluating interventions for pediatric perianal fistulizing CD included a clinically evident fistula tract, plus confirmation of radiologic disease activity, defined as at least one fistula or abscess on pelvic MRI. Including both radiologic and clinical outcomes as co-primary endpoints were considered appropriate by the panel, as was the inclusion of a patient reported outcome. All authors considered it inappropriate for a clinical trial to not have radiologic confirmation of a fistula at inclusion. Panelists were uncertain whether the primary endpoint for a pediatric perianal fistulizing CD study should be based upon a radiologic outcome alone.

Appropriate minimum technical requirements for radiologic assessment included imaging across three planes (sagittal, coronal and axial – the latter two aligned relative to the anal canal) and slice

thickness of ≤ 5 mm. There was some uncertainty regarding the field of view given the variation in pediatric patient sizes. The use of diffusion weighted imaging (DWI) was considered helpful when the quality of the scan is optimal, however that this sequence can be subject to artifact was viewed as a technical limitation. Additionally, the panel was uncertain about the use of DWI in clinical trials given the lack of data on the additional benefit of DWI over standard T2-weighted imaging and problems with reproducibility in a multisite clinical trial setting. Although highlighted as a particular concern for the patient population, the use of intravenous gadolinium-based contrast material for adequate assessment of perianal fistulizing CD activity in pediatric clinical trials was considered appropriate at present, and useful for distinguishing collections from inflammatory masses and purulent fluid from granulation tissue.

Panel uncertainty was found regarding the statement pertaining to adequacy of technical parameters when assessing fluid collections and fluid filled tracts, i.e., spatial resolution, signal-to-noise ratio, and echo time, and likely reflects differing technical parameters in use based on patient and site-specific preference. The panel was also uncertain regarding the use of spline (non-linear) tools for the measurement of fistula length, and concern was raised about the availability of this tool for the purpose of outcome measurement.

## DISCUSSION

There are few trials published evaluating the efficacy of medical or surgical interventions for pediatric perianal fistulizing CD. Clinical assessment alone may be inadequate and adult studies have evolved to use MRI-based imaging in combination with clinical assessment as the gold standard for evaluating perianal fistulizing disease activity. This circumstance has been possible due to the development of MRI-based indices for assessment. Specifically, MRI-based indices

have evolved to provide greater consistency in assessment of perianal disease extent and activity and are sufficiently responsive to enable their use in clinical practice and in clinical trials guiding drug development.[10, 11] Important and much needed research to aid development of MRI indices for assessment of perianal fistulizing CD in the pediatric population has only recently begun.[12]

In the first stage of this study, evaluating the operating properties of MRI indices for pediatric perianal fistulizing CD, we assessed the appropriateness of MRI-based fistulizing CD activity indices and their component items through a RAND/UCLA appropriateness process that included disease experts. MRI-based assessment of perianal fistulizing CD activity in pediatric clinical trials was considered appropriate by the RAND panel, and the mVAI and MAGNIFI-CD specifically were considered appropriate indices to assess radiologic disease activity. Panelists were not specifically asked about the appropriateness of the PEMPAC index, as this index was not yet available in the public domain when this RAND exercise was conducted. However, numerous items included in PEMPAC, such as number of fistulas, hyperintensity on T2-weighted images, location, and collections were considered appropriate by the panel. PEMPAC is a pediatric-specific, imaging-based index that was developed and validated using data from 88 children included in the ImageKids study.[12] Strong correlation of PEMPAC with the VAI ($r=0.92$, $p<0.001$) supports the validity of the instrument, but responsiveness data is limited, and our study addresses many aspects of a pediatric perianal clinical trial that are beyond an imaging-based index employed by central readers, e.g., technical parameters and requirements, inclusion criteria, clinical outcomes and patient-reporting outcome tools[12].

These results highlight the findings of a RAND appropriateness process involving a panel of 15 experts, 2 rounds of survey responses, and 2 moderated panel meetings. There was no panel disagreement regarding the appropriateness of a total of 79 statements. The only statement

considered inappropriate proposed a requirement for radiologic confirmation of a clinically active fistula as a criterion for inclusion of patients in clinical trials in pediatric perianal fistulizing CD. Number of fistula tracts, the predominant fistula location and length, complexity of the fistula, and fistula extension were considered appropriate items for the assessment of perianal fistulizing CD activity. Fistula length, which is not included as an item in the VAI or mVAI, is included in MAGNIFI-CD and PEMPAC and can have prognostic implications.[9-11] Shenoy-Bhangle et al[20] demonstrated absolute and aggregate fistula length on pelvic MRI were the best predictors of treatment response in a pediatric CD cohort, with 63% of patients with a dominant fistula length <2.5 cm demonstrating a good response (defined as a decrease in fistula length and abscess volume and no evidence of a new fistula or abscess) compared to 25% of patients with dominant fistula length ≥2.5 cm ($p < 0.05$). Additionally, in 42% of those studied, an aggregate or total fistula length ≥2.5 cm was associated with disease progression, compared to 8% of patients with aggregate length <2.5 cm. The RAND panel was uncertain whether sinus tracts should be differentiated and assessed separately from fistulas in MR images, as they may be difficult to distinguish when small or short in length. There was further debate on the severity rating of fistula extensions, with panelists disagreeing on the severity of horseshoe extensions relative to infralevatoric or supralevatoric extensions.

Assessment of hyperintensity of the primary tract on both fat-saturated T2-weighted images and post-contrast T1-weighted images as a component of perianal fistulizing activity were considered appropriate by the panel. Use of intravenous gadolinium-based contrast material was considered appropriate for adequate radiologic assessment of perianal fistulizing CD activity. The utility of DWI in the clinical trial setting, as well as in clinical practice, was considered unknown at the current time due to lack of data on its utility, as well as technical complexity and variability for

comparable purposes. This contrasts with considerations made during the development of the PEMPAC index. Although DWI features are not scored as a separate index item, DWI was used in conjunction with T2-weighted imaging to help delineate fistulas and abscesses in preference to contrast-enhanced T1-weighted imaging given the "Black box" warning regarding the use of gadolinium-based contrast agents. Scores for a subset of patients (7) were found to be similar when defining index items with and without use of contrast-enhanced images,[21] but experts on our panel considered evidence for reliable identification of abscesses without gadolinium is lacking, potentially compromising assessment of novel therapeutic agents. The panelists strongly felt external validation with larger numbers is required to clarify the contribution of DWI added to T2-weighted sequences in differentiating a collection from an inflammatory mass and its role in potentially replacing post-contrast fat-suppressed T1-weighted imaging.

Regarding the configuration of endpoints for clinical trials, a co-primary endpoint of a radiological and clinical outcome was considered appropriate by the panelists. Inclusion of a patient reported outcome measure was also appropriate.

Several limitations of this study should be acknowledged. Whilst we did aim for a multi-specialty panel, only two surgeons participated. Although we undertook subgroup analysis of subspecialty ratings of appropriateness (radiology versus gastroenterology and surgery combined) for uncertain items, the interpretation of these results is limited by small subgroup sample sizes. The results of this analysis are provided as a descriptive summary only (Supplemental Table 5). We did not specifically assess the appropriateness of the overall PEMPAC index, the reliability of this index will be assessed in an upcoming study. Our study also has limitations relating to available evidence and issues around uncertainty of specific statements, however these topics represent areas for which consensus or evidence is needed.

In conclusion, this study assessed the appropriateness of MRI-based disease activity indices, assessment methods, and clinical trial design for pediatric perianal fistulizing CD. This work provides a framework for the future assessment of the operating properties of existing MRI indices and items in a pediatric IBD cohort. A list of MRI items considered appropriate for assessing pediatric perianal fistulizing CD based on face validity, and that reliably measure disease activity will be collated. This is a crucial step in the development of a fully validated index. This would allow objective assessment of novel therapies to promote complete disease remission, thereby improving the quality of life of patients, and reducing the risk of relapse and caregiver burden.

**Table 1**. Summary of index items under consideration

| Item | VAI | mVAI | MAGNIFI-CD | PEMPAC |
|---|---|---|---|---|
| Number of fistula tracts | + | | + | + |
| Location | + | | | + |
| Extension | + | + | + | |
| Hyperintensity on T2-weighted images | + | + | | + |
| Hyperintensity on post-contrast T1-weighted images | | | + | |
| Collections | + | | | + |
| Rectal wall involvement/proctitis | + | + | | |
| Inflammatory mass | | + | + | |
| Dominant feature | | + | + | |
| Fistula length | | | + | |
| Aggregate length of fistulas | | | | + |

VAI, Van Assche Index; mVAI, modified Van Assche Index; MAGNIFI-CD, MAGnetic resonance Novel Index for Fistula Imaging in Crohn's Disease; PEMPAC, Pediatric MRI-based Perianal Crohn's disease index.

**Table 2**. Summary of RAND 2 survey results

| Item | Median, panel score | MAD | Appropriateness |
|---|---|---|---|
| (2.1) The Van Assche Index (VAI) should be used for assessment of radiological disease activity in pediatric perianal fistulizing CD clinical trials. | 5 | 1.40 | Uncertain |
| (3.1) The modified VAI (mVAI) should be used for assessment of radiological disease activity in pediatric perianal fistulizing CD clinical trials. | 7 | 1.20 | Appropriate |
| (4.1) The Magnetic Resonance Novel Index for Fistula Imaging in CD (MAGNIFI-CD) should be used for assessment of radiological disease activity in pediatric perianal fistulizing CD clinical trials. | 8 | 0.80 | Appropriate |
| (5.1) The St. James' University Hospital Classification should be used for assessment of radiological disease activity in pediatric perianal fistulizing CD clinical trials. | 5 | 1.40 | Uncertain |
| (6.1) The Parks Classification should be used for assessment of radiological disease activity in pediatric perianal fistulizing CD clinical trials. | 5 | 0.73 | Uncertain |
| (7.1) The number of fistula tracts should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 8 | 0.73 | Appropriate |
| (8.1) The predominant fistula location should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 8 | 1.33 | Appropriate |
| (9.1) The complexity of fistula tracts should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 8 | 0.60 | Appropriate |
| (10.1) Fistula extension should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 8 | 0.67 | Appropriate |
| (11.1) Sinuses (cutaneous opening) should be differentiated and assessed separately from fistulas in MR images obtained during pediatric perianal fistulizing CD trials. | 5 | 2.00 | Uncertain |
| (12.1) Hyperintensity of the primary tract or extensions (most severe lesion compared to signal intensity of nearby *in-plane vessels*) on fat-saturated T2-weighted images should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 8 | 0.87 | Appropriate |
| (13.1) Hyperintensity of the primary tract or extensions (most severe lesion compared to signal intensity of nearby *in-plane vessels*) on post-contrast fat-saturated T1-weighted images should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 7 | 1.00 | Appropriate |
| (14.1) The dominant feature (fibrous, granulation tissue, pus) of the primary tract and extensions should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 8 | 1.13 | Appropriate |
| (15.1) The length of the (predominant) fistula tract should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 7 | 1.27 | Appropriate |
| (16.1) Maximum fistula length (the longest axis diameter based on fistula evaluation in three planes) should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 7 | 1.53 | Appropriate |
| (17.1) Aggregate fistula length (the sum of maximum fistula lengths for all fistulas present) should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 5 | 1.33 | Uncertain |

| | | | |
|---|---|---|---|
| (18.1) The number of internal openings should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 5 | 1.67 | Uncertain |
| (19.1) The number of external openings should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 7 | 1.40 | Appropriate |
| (20.1) The presence of a recto/anovaginal tract should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 7 | 1.93 | Appropriate |
| (21.1) The presence of collections (cavities > 3 mm in diameter) should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 8 | 0.60 | Appropriate |
| (22.1) The presence of abscesses (rim enhancing, fat-saturated T2-weighted hyperintense/post-contrast fat-saturated T1-weighted hypointense contents and/or T1- and T2-weighted hypointense gas and/or restricting [if diffusion-weighted images]) should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 9 | 0.73 | Appropriate |
| (23.1) Abscess volume should be measured as a component of perianal fistulizing CD activity in pediatric clinical trials. | 7 | 0.93 | Appropriate |
| (24.1) The presence of an inflammatory mass should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 8 | 0.73 | Appropriate |
| (25.1) Rectal wall involvement should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 8 | 1.27 | Appropriate |
| (26.1) Extra-anal regional inflammation (small foci of inflammation near natal cleft or subcutaneous fat in buttock not obviously connecting to anal canal) should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 6 | 1.40 | Uncertain |
| (68.0) Hyperintensity of the primary tract or extensions (most severe lesion compared to signal intensity of *nearby fluid* on fat-saturated T2-weighted images) should be assessed as a component of perianal fistulizing CD activity in pediatric clinical trials. | 7 | 0.73 | Appropriate |
| (27.1) The number of fistula tracts should be scored as<br>0 = none<br>1 = single, unbranched<br>2 = single, branched 3 = multiple | 7 | 1.80 | Appropriate |
| (28.1) The number of fistula tracts should be scored as<br>0 = none<br>1 = single, unbranched<br>2 = complex (includes single, branched, or multiple tracts) | 8 | 1.13 | Appropriate |
| (29.1) The number of fistula tracts should be scored as<br>0 = none (no tracts visible)<br>1 = single, unbranched (single internal opening leading to a single fistula tract [internal opening defined as discontinuation of anal mucosa or closest proximity of tract to anal mucosa])<br>2 = complex (either a single internal opening leading to more than 1 fistula tract or multiple internal openings) | 7 | 1.33 | Appropriate |
| (30.1) The location of the primary tract should be scored as<br>1 = extra- or intersphincteric<br>2 = transsphincteric<br>3 = suprasphincteric | 7 | 1.60 | Appropriate |

| | | | |
|---|---|---|---|
| (31.1) The location of the most dominant feature should be scored as<br>0 = submucosal (tract lies superficial to the internal sphincter)<br>1 = intersphincteric (tract extends through the internal sphincter to the intersphincteric plane then to the perineal skin)<br>2 = transsphincteric (tract extends via the internal and external anal sphincter [or puborectalis muscle] into the ischioanal fossa then to the perineal skin)<br>3 = extrasphincteric (tract extends through the ischioanal fossa, upwards and through the levator ani muscles to the rectal wall completely outside the sphincter mechanism)<br>4 = suprasphincteric (tract extends via the intersphincteric space, then tracts superiorly to above the puborectalis muscle [i.e., above the anorectal junction] before curving downward through the levator muscle lateral to the external anal sphincter and puborectalis muscle into the ischioanal fossa then to the perineal skin) | 8 | 1.33 | Appropriate |
| (32.1) The complexity of perianal fistulas should be scored as<br>0 = normal appearance<br>1 = simple linear intersphincteric fistula<br>2 = intersphincteric fistula with intersphincteric abscess or secondary fistulous track<br>3 = transsphincteric fistula<br>4 = transsphincteric fistula with abscess or secondary track within the ischioanal or ischiorectal fossa<br>5 = supralevator and translevator disease | 8 | 1.80 | Appropriate |
| (33.1) Extension of the primary fistula tract should be scored as 1 = infralevatoric 2 = supralevatoric | 7 | 1.27 | Appropriate |
| (34.1) When assessing fistula extension, all relevant findings should be identified, and the highest score chosen from the following<br>0 = absent (no extension)<br>1 = infralevatoric (extends upward in the ischioanal fossa but remains below the levator ani muscle)<br>2 = horseshoe configuration (extends into the intersphincteric space on both sides of the midline)<br>3 = supralevatoric (any extension in the supralevatoric space [i.e., above where the levator plate is connected to the anorectum]) | 7 | 1.07 | Appropriate |
| (35.1) When assessing fistula extension, the most severe finding should be scored as<br>0 = absent (no extension)<br>1 = horseshoe (extends into the intersphincteric space on both sides of the midline)<br>2 = infralevatoric (extends upward in the ischioanal fossa but remains below the levator ani muscle) or supralevatoric (any extension in the supralevatoric space [i.e., above where the levator plate is connected to the anorectum]) | 6 | 1.67 | Uncertain |
| (36.1) Sinuses should be scored as<br>0 = absent<br>1 = present | 7 | 1.13 | Appropriate |
| (37.1) Hyperintensity (of the primary fistula tract) on fat-saturated T2-weighted images should be scored as<br>0 = absent<br>4 = mild<br>8 = pronounced | 7 | 1.00 | Appropriate |
| (38.1) Hyperintensity (of the primary fistula tract or extension) on fat-saturated T2-weighted images should be scored as<br>0 = absent (no hyperintensity visible, only scar tissue)<br>1 = mild (slight increase in signal intensity but less than nearby in-plane vessels)<br>2 = pronounced (tract showing equal or greater signal hyperintensity than nearby in-plane vessels) | 8 | 0.80 | Appropriate |

| | | | |
|---|---|---|---|
| (39.1) Hyperintensity of (the most severe) primary tract or extensions on post-contrast fat-saturated T1-weighted images should be scored as<br>0 = absent (no hyperintensity visible)<br>1 = mild (slight increase in signal intensity but less than nearby in-plane vessels)<br>2 = pronounced (tract showing equal or greater signal hyperintensity than nearby in-plane vessels) | 7 | 0.80 | Appropriate |
| (40.1) Hyperintensity of (the most severe) primary tract or extensions on post-contrast fat-saturated T1-weighted images should be scored as<br>0 = no hyperintensity visible or slight increase in signal intensity but less than nearby in-plane vessels<br>1 = tract showing equal or greater signal hyperintensity than nearby in-plane vessels | 7 | 1.47 | Appropriate |
| (41.1) The dominant feature of primary tract and extensions should be scored as<br>0 = predominantly fibrous (> 50% of tract has a fibrotic appearance [i.e., hypointense on fat-saturated T2-weighted images])<br>1 = predominantly filled with granulation tissue (> 50% of tract is filled with granulation tissue [i.e., hyperintense on fat-saturated T2-weighted images with enhancement of contents and wall on post-contrast fat-saturated T1-weighted images])<br>2 = predominantly filled with fluid or pus (> 50% of tract is filled with fluid or pus [i.e., hyperintense on fat-saturated T2-weighted images with no enhancement of contents on post-contrast fat-saturated T1-weighted images, though lining of tract may enhance]) | 8 | 1.00 | Appropriate |
| (42.1) Length of the fistula tract should be scored as<br>0 = less than 2.5 cm<br>1 = 2.5 to 5.0 cm<br>2 = greater than 5 cm | 7 | 1.00 | Appropriate |
| (43.1) The aggregate length of all fistula tracts should be scored as<br>0 = < 2.5 cm<br>1 = ≥ 2.5 cm | 7 | 1.20 | Appropriate |
| (44.1) The presence of recto/anovaginal tract should be scored as<br>0 = absent (no recto/anovaginal tract)<br>1 = rectovaginal tract (fistula arises from rectal mucosa)<br>2 = anovaginal tract (fistula arises from anal mucosa) | 7 | 1.20 | Appropriate |
| (45.1) Number of internal openings should be scored as<br>0 = 0<br>1 = 1<br>2 = 2<br>3 = greater than 2 | 7 | 1.27 | Appropriate |
| (46.1) Number of external openings should be scored as<br>0 = 0<br>1 = 1<br>2 = 2<br>3 = 3<br>4 = greater than 3 | 7 | 1.40 | Appropriate |
| (47.1) Collections (cavities > 3 mm in diameter) should be scored as<br>0 = absent<br>4 = present | 7 | 1.13 | Appropriate |
| (48.1) Abscesses (rim enhancing, fat-saturated T2-weighted hyperintense/ post-contrast fat saturated T1-weighted hypointense contents and/or T1- and T2-weighted hypointense gas and/or restricting [if diffusion weighted images]) should be assessed as<br>0 = absent<br>1 = present | 8 | 1.27 | Appropriate |

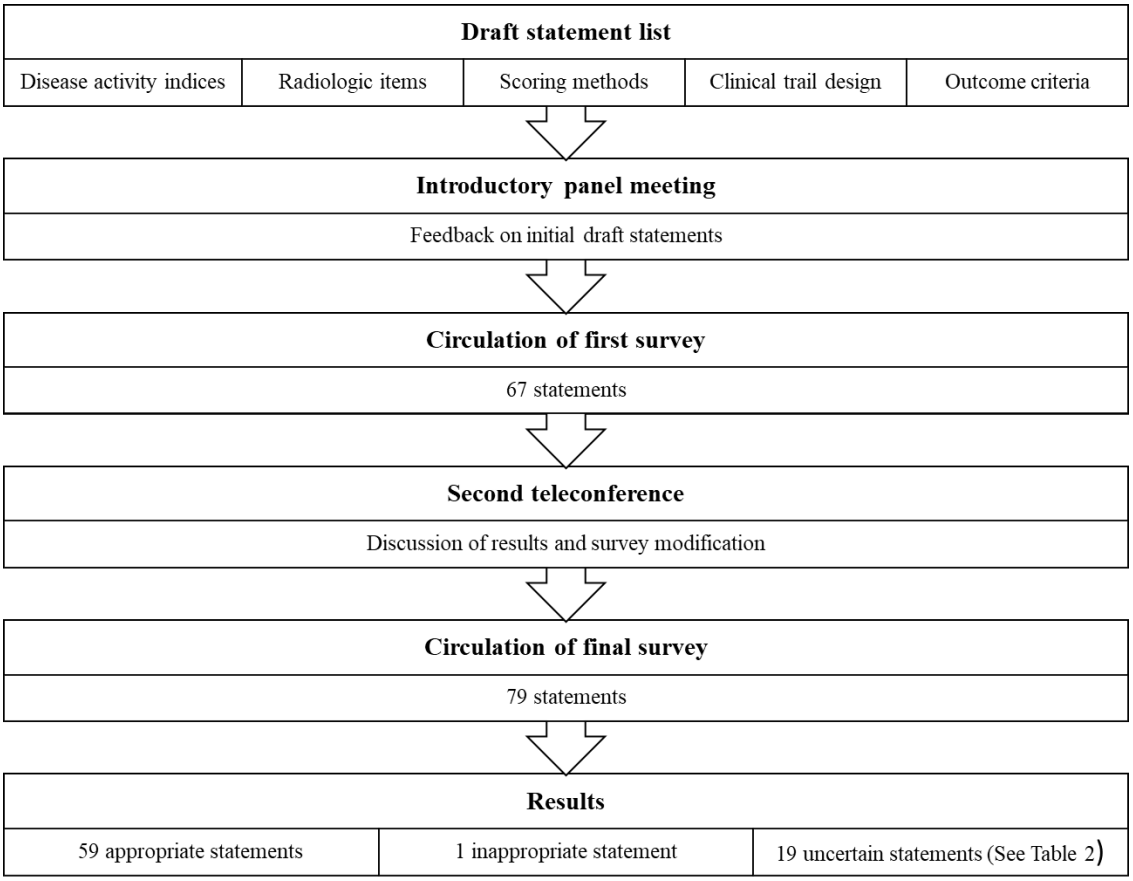
| | | | |
|---|---|---|---|
| (49.1) Abscess volume should be measured according to the ellipsoid volume formula (0.52 * long-axis diameter * [short axis diameter]2). | 5 | 1.53 | Uncertain |
| (50.1) Assessment of inflammatory mass should be performed by identifying all relevant findings, measuring fluid collections (excluding the wall) on the shortest axis, and scored as<br>0 = absent (no inflammatory mass)<br>1 = diffuse (diffuse inflammation of surrounding tissues)<br>2 = focal (lesion > 3 mm in diameter on fat-saturated T2-weighted images [but does not include linear tracts with diameter > 3 mm] with diffuse enhancement on post-contrast fat- saturated T1-weighted images [i.e., granulation tissue])<br>3 = collection–small (circumscribed cavity 3-10 mm in diameter [but does not include linear tracts with diameter > 3 mm and if present they should be excluded from the measurement of the infiltrate]). Hyperintense appearance on fat-saturated T2-weighted images with rim enhancement on post-contrast fat-saturated T1-weighted images<br>4 = collection–medium (as defined above except diameter measures 11-20 mm)<br>5 = collection–large (as defined above except diameter measures > 20 mm) | 7 | 1.47 | Appropriate |
| (51.1) For assessment of the inflammatory mass item, an infiltrate should be described as an inflammatory region as a whole that may or may not include one or more fluid collections. | 5 | 1.40 | Uncertain |
| (52.1) For assessment of the inflammatory mass item, all relevant findings should be identified, and the highest score chosen. | 7 | 0.60 | Appropriate |
| (53.1) For assessment of the inflammatory mass item, fluid collections should be measured on the shortest axis and exclude the wall. | 7 | 1.67 | Appropriate |
| (54.1) Rectal wall involvement should be scored as<br>0 = normal<br>2 = thickened | 6 | 1.53 | Uncertain |
| (55.1) Rectal wall involvement should be scored as<br>0 = normal (normal appearance of rectal wall)<br>1 = thickened (thickened rectal wall [>3 mm when distended])<br>2 = increased signal intensity (hyperintensity of the rectal wall on fat-saturated T2-weighted images [compared to nearby in-plane vessels], mural stratification and/or perimural infiltrate) | 7 | 1.47 | Appropriate |
| (56.1) Proctitis should be scored as<br>0 = absent (normal appearance of rectal wall)<br>2 = present (increased wall thickness and size of mesorectal lymph nodes [> 5 mm], creeping fat, increased perimural T2 signal and enhancement) | 7 | 1.40 | Appropriate |
| (69.0) When assessing fistula extension, all relevant findings should be identified, and the highest score chosen from the following<br>0 = extension absent (inter/transsphincteric fistula without coursing upward/downward/across midline)<br>1 = infralevatoric extension (inter/transsphincteric coursing upward/downward but below levator ani)<br>2 = horseshoe configuration (inter/transsphincteric fistula either side of internal opening, 2 or more quadrants)<br>3 = complex extension (inter/transsphincteric fistula with infralevator extension AND horseshoe configuration)<br>4 = supralevator extension (inter/transsphincteric fistula with any extension above levator ani) | 7 | 0.93 | Appropriate |
| (70.0) A horseshoe extension is defined as one that extends into the intersphincteric space on both sides of the midline. | 7 | 1.47 | Appropriate |

| | | | |
|---|---|---|---|
| (71.0) A horseshoe extension configuration is a semilunar region of sepsis that spreads in the horizontal plane either side of an internal opening, to involve two or more adjacent quadrants. Horseshoe extensions may be ischioanal, intersphincteric, or supralevator. | 8 | 0.60 | Appropriate |
| (72.0) Hyperintensity (of the primary fistula tract or extension) on fat-saturated T2-weighted images should be scored as<br>0 = absent (no hyperintensity visible, only scar tissue)<br>1 = mild (slight increase in signal intensity but less than nearby fluid)<br>2 = pronounced (tract showing equal or greater signal hyperintensity than nearby fluid) | 8 | 1.07 | Appropriate |
| (73.0) Assessment of inflammatory mass should be performed by identifying all relevant findings, measuring fluid collections (excluding the wall) on the shortest axis, and scored as<br>0 = absent (no inflammatory mass)<br>1 = focal (lesion > 3 mm in diameter on fat-saturated T2-weighted images [but does not include linear tracts with diameter > 3 mm] with diffuse enhancement on post-contrast fat-saturated T1-weighted images [i.e., granulation tissue])2 = diffuse (diffuse inflammation of surrounding tissues)<br>3 = collection–small (circumscribed cavity 3-10 mm in diameter [but does not include linear tracts with diameter > 3 mm and if present they should be excluded from the measurement of the infiltrate]). Hyperintense appearance on fat-saturated T2-weighted images with rim enhancement on post-contrast fat-saturated T1-weighted images<br>4 = collection–medium (as defined above except diameter measures 11-20 mm)<br>5 = collection–large (as defined above except diameter measures > 20 mm) | 7 | 0.87 | Appropriate |
| (74.0) Rectal wall involvement should be scored as<br>0 = normal (normal appearance of rectal wall)<br>1 = thickened (thickened rectal wall [>3 mm when distended])<br>2 = increased signal intensity (hyperintensity of the rectal wall on fat-saturated T2-weighted images [compared to nearby fluid], mural stratification and/or perimural infiltrate). | 8 | 1.33 | Appropriate |
| (57.1) Inclusion criteria for clinical trials in pediatric perianal fistulizing CD should require a clinically evident fistula tract, plus confirmation of radiological disease activity defined as at least one fistula or abscess on pelvic MRI. | 8 | 1.60 | Appropriate |
| (58.1) Inclusion criteria for clinical trials in pediatric perianal fistulizing CD should not require radiologic confirmation of clinically evident fistula. | 2 | 2.33 | Inappropriate |
| (59.1) The primary endpoint of pediatric perianal fistulizing CD clinical trials should be based on a radiological outcome. | 6 | 1.60 | Uncertain |
| (60.1) Pediatric perianal fistulizing CD clinical trials should include a co-primary endpoint of a radiological and clinical outcome. | 8 | 0.80 | Appropriate |
| (61.1) Assessment of radiological disease activity should be a secondary endpoint in pediatric perianal fistulizing CD clinical trials. | 4 | 1.53 | Uncertain |
| (62.1) Pediatric perianal fistulizing CD clinical trials should include a patient reported outcome. | 8 | 0.80 | Appropriate |
| (63.1) Diffusion weighted imaging should be included as a component of radiological assessment of perianal fistulizing CD activity in pediatric clinical trials. | 6 | 1.60 | Uncertain |
| (64.1) Radiologic assessment of pediatric perianal fistulizing CD activity should include items that are responsive to change with a seton in place. | 8 | 1.27 | Appropriate |

| | | | |
|---|---|---|---|
| (65.1) Perianal fistulizing CD activity can be adequately measured radiologically without the use of contrast to assess outcomes in pediatric clinical trials. | 5 | 0.67 | Uncertain |
| (66.1) Spatial resolution, signal to noise ratio, and echo time are adequate to define true fluid collections and fluid-filled tracts. | 5 | 0.93 | Uncertain |
| (67.1) Spline tools can accurately measure the length of a fistula tract. | 5 | 0.93 | Uncertain |
| (75.0) At present, the use of contrast is required for adequate radiologic assessment of perianal fistulizing CD activity in pediatric clinical trials. | 7 | 1.33 | Appropriate |
| (76.0) Minimum technical requirements for radiologic assessment of perianal fistulizing CD activity in pediatric clinical trials include | | | |
| (76.0.a) Three imaging planes – sagittal, coronal (parallel to anal canal) and axial (orthogonal to anal canal). | 8 | 1.20 | Appropriate |
| (76.0.b) Slice thickness ≤ 5 mm | 8 | 1.47 | Appropriate |
| (76.0.c) Small field of view (FOV) with 1 larger FOV sequence for more remote extensions | 5 | 1.60 | Uncertain |
| (76.0.d) Diffusion weighted imaging optional | 5 | 1.40 | Uncertain |

MAD, mean absolute deviation; MR magnetic resonance; CD Crohn's disease; mm, millimeters.

**Figure 1**. Statement generation and survey



## REFERENCES


1. Dhaliwal J, Walters TD, Mack DR, et al. Phenotypic Variation in Paediatric Inflammatory Bowel Disease by Age: A Multicentre Prospective Inception Cohort Study of the Canadian Children IBD Network. J Crohns Colitis 2020;14:445-454.
2. van Rheenen PF, Aloi M, Assa A, et al. The Medical Management of Paediatric Crohn's Disease: an ECCO-ESPGHAN Guideline Update. J Crohns Colitis 2020.
3. de Zoeten EF, Pasternak BA, Mattei P, et al. Diagnosis and treatment of perianal Crohn disease: NASPGHAN clinical report and consensus statement. J Pediatr Gastroenterol Nutr 2013;57:401-12.
4. Forsdick VK, Tan Tanny SP, King SK. Medical and surgical management of pediatric perianal Crohn's disease: A systematic review. J Pediatr Surg 2019;54:2554-2558.
5. Lee T, Yong E, Ding NS. Radiological outcomes in perianal fistulizing Crohn's disease: A systematic review and meta-analysis. JGH Open 2020;4:340-344.
6. AlSabban Z, Carman N, Moineddin R, et al. Can MR enterography screen for perianal disease in pediatric inflammatory bowel disease? J Magn Reson Imaging 2018;47:1638-1645.
7. Parks AG, Gordon PH, Hardcastle JD. A classification of fistula-in-ano. Br J Surg 1976;63:1-12.
8. Spencer JA, Ward J, Beckingham IJ, et al. Dynamic contrast-enhanced MR imaging of perianal fistulas. AJR Am J Roentgenol 1996;167:735-41.
9. Van Assche G, Vanbeckevoort D, Bielen D, et al. Magnetic resonance imaging of the effects of infliximab on perianal fistulizing Crohn's disease. Am J Gastroenterol 2003;98:332-9.
10. Samaan MA, Puylaert CAJ, Levesque BG, et al. The development of a magnetic resonance imaging index for fistulising Crohn's disease. Aliment Pharmacol Ther 2017;46:516-528.
11. Hindryckx P, Jairath V, Zou G, et al. Development and Validation of a Magnetic Resonance Index for Assessing Fistulas in Patients With Crohn's Disease. Gastroenterology 2019;157:1233-1244 e5.
12. Choshen S, Turner D, Pratt LT, et al. Development and Validation of a Pediatric MRI-Based Perianal Crohn Disease (PEMPAC) Index-A Report from the ImageKids Study. Inflamm Bowel Dis 2021.
13. Greer MC, Taylor SA. Perianal Imaging in Crohn Disease: Current Status With a Focus on MRI, From the AJR Special Series on Imaging of Inflammation. AJR Am J Roentgenol 2022:1-12.
14. Fitch K, Bernstein SJ, Aguilar MD, et al. The RAND/UCLA Appropriateness Method User's Manual. Santa Monica, CA: RAND Corporation, 2001.
15. Novak G, Parker CE, Pai RK, et al. Histologic scoring indices for evaluation of disease activity in Crohn's disease. Cochrane Database Syst Rev 2017;7:Cd012351.
16. Brook RH. The RAND/UCLA appropriateness method. . In: McCormick KA, Moore SR, Siegel RA, eds. Clinical practice guideline development: Methodology practices (AHCPR Publication No. 95-0009). Rockville, MD: Agency for Health Care Policy and Research, Public Health Service, U.S. Department of Health and Human Services, 1994.
17. Fitch K BS, Aguilar M, Burnand B, LaCalle J, Lazaro P, et al. The RAND/UCLA Appropriateness Method User’s Manual. Santa Monica, California: RAND Corporation, 2001.
18. Sheedy SP, Bruining DH, Dozois EJ, et al. MR Imaging of Perianal Crohn Disease. Radiology 2017;282:628-645.
19. Halligan S, Tolan D, Amitai MM, et al. ESGAR consensus statement on the imaging of fistula-in-ano and other causes of anal sepsis. Eur Radiol 2020;30:4734-4740.
20. Shenoy-Bhangle A, Nimkin K, Goldner D, et al. MRI predictors of treatment response for perianal fistulizing Crohn disease in children and young adults. Pediatr Radiol 2014;44:23-9.
21. Administration FD. FDA Drug Safety Communication: FDA warns that gadolinium-based contrast agents (GBCAs) are retained in the body; requires new class warnings, 2018.

## LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

| | |
|---|---|
| CD | Crohn's disease |
| DWI | diffusion-weighted imaging |
| IBD | inflammatory bowel disease |
| MAD | mean absolute deviation |
| MAGNIFI-CD | MAGnetic resonance Novel Index for Fistula Imaging in Crohn's Disease |
| MRI | magnetic resonance imaging |
| mVAI | modified Van Assche Index |
| PEMPAC | Pediatric MRI-based Perianal Crohn's disease |
| RAND/UCLA | Research and Development/University of California, Los Angeles |
| TNF | tumor necrosis factor |
| VAI | Van Assche Index |

## AUTHORSHIP CONTRIBUTIONS:

Study concept and design: EC, CM, BGF, AMG, MLG, VJ

Acquisition of data: EC, LG, JR, LS, LMS

Analysis and interpretation of data: EC, LG, GZ, CM, LMS, MLG, VJ, AMG

Drafting of the manuscript: EC, VJ

Critical revision of the manuscript for important intellectual content: all authors