



Published in final edited form as:

Gastroenterology. 2020 January ; 158(1): 137–150.e1. doi:10.1053/j.gastro.2019.08.033.

Histopathology Scoring Systems of Stenosis Associated With Small Bowel Crohn's Disease: A Systematic Review

Ilyssa O. Gordon^{*1}, Dominik Bettenworth^{*2}, Arne Bokemeyer², Amitabh Srivastava³, Christophe Rosty⁴, Gert e Hertogh⁵, Marie E. Robert⁶, Mark A. Valasek⁷, Ren Mao^{8,9}, Satya Kurada^{9,10}, Noam Harpaz¹¹, Paula Borralho¹², Reetesh K. Pai¹³, Rish K. Pai¹⁴, Robert Odze¹⁵, Roger Feakins¹⁶, Claire E. Parker¹⁷, Tran Nguyen¹⁷, Vipul Jairath^{17,18,19}, Mark E. Baker²⁰, David H. Bruining²¹, J.G. Fletcher²², Brian G. Feagan^{17,18,19}, Florian Rieder^{9,10}
Stenosis Therapy and Anti-Fibrotic Research (STAR) Consortium

¹Department of Pathology, Robert J. Tomsich Pathology and Laboratory Medicine Institute, Cleveland Clinic Foundation, Cleveland, USA

Correspondence: Florian Rieder M.D., Department of Gastroenterology, Hepatology & Nutrition, Digestive Diseases and Surgery Institute, 9500 Euclid Avenue, Cleveland, OH, 44195, USA; Phone: +1 (216) 445 4916, Fax: +1 (216) 636 0104; riederf@ccf.org.

*these authors contributed equally and share first authorship

AUTHORS CONTRIBUTION

Ilyssa Gordon: study concept and design; acquisition of data; drafting of the manuscript; critical revision of the manuscript for intellectual content; analysis and interpretation of data.

Dominik Bettenworth: study concept and design; acquisition of data; drafting of the manuscript; critical revision of the manuscript for intellectual content; analysis and interpretation of data.

Arne Bokemeyer: acquisition of data; analysis and interpretation of data; critical revision of the manuscript for intellectual content; administrative and technical support.

Amitabh Srivastava: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Christophe Rosty: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Gert De Hertogh: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Marie Robert: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Mark A. Valasek: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Ren Mao: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Satya Kurada: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Noam Harpaz: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Paula Borralho: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Reetesh K. Pai: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Rish K. Pai: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Robert Odze: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Roger Feakins: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Claire E. Parker: study concept and design; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content; administrative, technical and material support

Tran Nguyen: study concept and design; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content; administrative, technical and material support

Vipul Jairath: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Mark Baker: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

David Bruining: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

JG Fletcher: analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content

Brian Feagan: analysis and interpretation of data; drafting of the manuscript; obtained funding; critical revision of the manuscript for intellectual content; study supervision.

Florian Rieder: study concept and design; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for intellectual content; administrative, technical and material support; study supervision.

Authors marked in **bold** contributed equally to the respective manuscript.

²Department of Medicine B, Gastroenterology and Hepatology, University Hospital Münster, Münster, Germany

³Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA

⁴Department of Clinical Pathology, The University of Melbourne, Parville, VIC, Australia; Envoi Specialist Pathologists, Brisbane, QLD, Australia; Faculty of Medicine, University of Queensland, Brisbane, QLD, Australia

⁵Department of Pathology, University Hospital Leuven, Leuven, Belgium

⁶Department of Pathology, Yale University School of Medicine, New Haven, USA

⁷Department of Pathology, University of California San Diego, La Jolla, California, USA

⁸Department of Gastroenterology, First Affiliated Hospital of Sun Yat-sen University, China

⁹Department of Inflammation and Immunity, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, Ohio, USA

¹⁰Department of Gastroenterology, Hepatology and Nutrition, Digestive Diseases and Surgery Institute; Cleveland Clinic Foundation, Cleveland, USA

¹¹Department of Pathology, Molecular and Cell-Based Medicine, Icahn School of Medicine at Mount Sinai, New York, NY

¹²Faculdade de Medicina da Universidade de Lisboa, Instituto de Anatomia Patológica, Lisbon, Portugal

¹³Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

¹⁴Department of Laboratory Medicine and Pathology, Mayo Clinic Arizona, Scottsdale, AZ, USA

¹⁵Pathology Department, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.

¹⁶Department of Cellular Pathology, Royal London Hospital, Barts Health NHS Trust, London, United Kingdom

¹⁷Robarts Clinical Trials, Inc., London, ON, Canada

¹⁸Department of Medicine, Western University, London, ON, Canada

¹⁹Department of Biostatistics and Epidemiology, Western University, London, ON, Canada

²⁰Department of Diagnostic Radiology, Imaging Institute, Digestive Diseases and Surgery Institute, and Taussig Cancer Institute, Cleveland Clinic Foundation, Cleveland, USA

²¹Division of Gastroenterology and Hepatology, Mayo Clinic College of Medicine, Rochester, Minnesota

²²Department of Radiology, Mayo Clinic, Rochester, USA

Abstract

Background & Aims: Stenosis is a common complication of Crohn's disease (CD) that has no effective medical therapy. Development of anti-fibrotic agents will require testing in randomized

controlled trials. Computed tomography enterography- and magnetic resonance enterography-based technologies might be used to measure outcomes in these trials. These approaches have been validated in studies of patients with symptomatic strictures who underwent imaging evaluations, followed by resection with histopathologic grading of the intestinal tissue for inflammation and/or fibrosis (the reference standard). Imaging findings have correlated with findings from quantitative or semi-quantitative histologic evaluation of the degree of fibromuscular stenosis and/or inflammation on the resection specimen. However, it is not clear whether histologic findings are an accurate reference standard. We performed a systematic review of all published histologic scoring systems used to assess stenosing CD.

Methods: We performed a comprehensive search of the Embase and Medline of studies through March 13, 2019 that used a histologic scoring system to characterize small bowel CD and assessed inflammatory and fibrotic alterations within the same adult subject. All scores fitting the criteria were included in our analysis, independently of the presence of stricturing disease, as long as inflammation and fibrosis were evaluated separately but in the same scoring system.

Results: We observed substantial heterogeneity among scoring systems, which were not derived using modern principles for evaluative index development. None had undergone formal validity or reliability testing. None of the existing indices had been constructed according to accepted methods for development of evaluative indices. Basic knowledge regarding their operating properties were lacking. Specific indices to evaluate the important pathological component of myofibroblast hypertrophy or hyperplasia have not been proposed.

Conclusions: In a systematic review of publications, we found a lack of validated histopathologic scoring systems for assessment of fibromuscular stenosis. Data that describe the operating properties of existing cross-sectional imaging techniques for stenosing CD should be questioned. Development and validation of a histopathology index is an important research priority.

Keywords

Fibrosis; Histopathology; IBD; Stricture

INTRODUCTION

Stenosing Crohn's disease (CD) ultimately affects more than one half of CD patients.^{1, 2} Although up to 20% of patients with small bowel stricturing disease are asymptomatic³, these individuals characteristically experience severely impaired quality of life. Progression of stenosis results in endoscopic^{4, 5} or surgical intervention,⁶ despite the use of anti-inflammatory therapies. The location of strictures follows the location of inflammation, with small bowel strictures being the most common site². Given this situation, considerable interest has evolved in the development of specific anti-fibrotic therapies for the disease. In recent years, progress has been made in developing novel therapies for several fibrosing disorders, including those involving the liver, lung, skin, kidney, and heart.⁷⁻¹¹ Based upon the introduction of specific anti-fibrotic agents in these conditions, controlled studies will soon be initiated in stenosing CD.

Success in clinical trials, however, will be predicated on development of validated endpoints. In this regard, cross-sectional imaging is likely to play a critical role in defining treatment efficacy.¹² Accordingly, multiple computed tomography enterography (CTE) and magnetic resonance enterography (MRE) imaging findings and measurements have been proposed as outcome measures for stenosing CD. Typically, candidate imaging modalities are evaluated in patients with symptomatic strictures, who undergo imaging, followed by bowel resection with grading of the intestinal tissue for inflammation and fibrosis.^{13–20} Construct validity is subsequently assessed by correlating imaging results to quantitative or semi quantitative histologic evaluations of the degree of fibromuscular stenosis and/or inflammation on the resection specimen. Several measurement challenges are inherent to this situation. First, CD-associated strictures are a transmural process²¹ characterized by overlap between inflammation and fibrosis including myofibroblast and myocyte hypertrophy and hyperplasia.^{14, 22–24} Multiple experimental imaging techniques have been tested to accurately separate inflammation from the other pathological components with the prospect of including them as outcome measures for future clinical trials,^{22, 25–27} however, none of these modalities have been validated according to contemporary methodological standards for assessment of validity, reliability and responsiveness.²⁸ Second, the correlational analyses necessary for establishing construct validity of cross sectional imaging modalities also require validated histologic indices for independently quantifying the relative amounts of inflammation and fibrosis present in the surgical specimen. Third, scoring of fibrosis and inflammation should be performed without knowledge of clinical information by more than one independent expert gastrointestinal pathologist whose intra-observer reliability for evaluating index items is well defined. The development and refinement of non-invasive, cross-sectional imaging methods to estimate the relative proportions and amount of inflammation, fibrosis and muscular hypertrophy within Crohn's disease strictures is dependent upon reproducible histopathologic validation. A reliable and validated histopathologic standard would provide discriminatory ability to improve and compare developing imaging techniques in computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound.

Based upon the conclusion that development of a validated histologic scoring system for stenosing CD is a rate limiting step to progress in this field, we established an international working group comprised of inflammatory bowel disease pathologists and gastroenterologists. This systematic review of all published histologic scoring systems for small bowel CD was performed as a first step towards the development of a fully validated histologic index for CD-associated small bowel strictures. We focused the review on small bowel strictures only, as this is the most common site for CD-associated strictures and is likely to be the first population to enter a clinical trial for an anti-fibrotic drug based on an international expert consensus²⁹.

METHODS

A formal systematic review with a comprehensive literature search, following PRISMA guidelines, was performed on March 13, 2019 to assess all relevant citations found in Embase and Medline (see Supplement Methods). Additionally, references of cited original articles and reviews were reviewed for identification of additional relevant publications.

After removing the duplicates from our search, all identified abstracts as well as full text publications of potentially eligible studies were assessed by both D. B. and A. B. Disagreement regarding inclusion/exclusion or extraction were resolved by F. R. In brief, the inclusion criteria for this systematic review were solely full thickness small bowel histopathology, which used a histologic scoring system to characterize small bowel CD assessing inflammatory as well as fibrotic alterations within the same subject. All scores fitting the criteria were included, independently of the presence of stricturing disease, as long as inflammation and fibrosis were evaluated separately, but in the same scoring system. Exclusion criteria were studies only available in abstract form, non-English language publications, patients younger than 18 years of age, studies containing less than 3 patients, review articles, colonic or upper GI strictures, ileal pouch anal anastomotic strictures, non-human subjects, full thickness histopathology not available in all patients of the study, histological evaluation of only inflammation alone *or* fibrosis alone, but not both, exclusive use of quantitative digital morphometry without histopathologic categorization, and no evaluation of CD. If the same score was used in several publications we only included the initial publication. The resulting studies were included in the qualitative analysis. The search included studies from the inception to March 2019. All other studies were excluded. A total of 13 studies satisfied these inclusion criteria (Table 1; Supplemental Figure 1).
14, 17, 22, 27, 30–38

RESULTS

Overview of Scoring Systems

A minority of the included studies exclusively assessed patients with stricturing disease phenotype (Supplemental Table 1).^{30, 34, 35} Half of the studies do^{14, 30, 36} or probably do^{34, 35, 37} assess both, the site of the stricture and the site of the adjacent non-strictured bowel from the same specimen (Supplemental Table 1). Most published scores included items in addition to inflammation and fibrosis, such as edema or muscle hypertrophy, 17, 22, 30, 32–36, 38 and half did not explicitly separate out an inflammatory score from a fibrosis score, but comment on both within the same score.^{30, 32–36} A summary of the different scoring systems can be found in Table 1.

Histopathology of Inflammation in Crohn's Disease Strictures

Histologic features of inflammation within a stricture overlap with those observed in non-strictured CD segments. The inflammatory elements include both active and chronic components (Figure 1).^{39–41} Activity is characterized by neutrophilic inflammation, ranging from cryptitis, to crypt abscess, erosion and ulcer. Chronic mucosal injury comprises both architectural distortion and chronic inflammation of the lamina propria. Paneth cell hyperplasia⁴² and pyloric gland metaplasia can also be present. The mucosal surface within a strictured segment is variable among patients, with some showing the typical chronic active disease pattern, some being completely ulcerated, and some showing chronic quiescent disease. Inflammation within the bowel wall includes fistula tracts, which are often associated with strictures, and transmural inflammation, which refers to lymphoid aggregates within the bowel wall and which are especially prominent at the interface between the muscularis propria and the mesenteric fat.

Additional inflammatory features which may be present in CD and CD-associated strictures include epithelioid granulomas and prominent eosinophils within the lamina propria.^{43, 44} Epithelioid granulomas are well-formed and non-necrotizing and can be found in any layer of the bowel wall. Microgranulomas, which are smaller clusters of epithelioid histiocytes not associated with ruptured crypts, are found in the mucosa.⁴⁵ Although not specifically studied to date in strictures, inflammation of submucosal and myenteric nerves and ganglia, or plexitis, has been identified in CD as predictors of post-operative recurrence.⁴⁶⁻⁴⁹

Histopathology of Fibrosis in Crohn's Disease Strictures

Our review of the literature identified that in CD, normal submucosal collagen and adipose tissue are replaced by excess extracellular matrix (ECM), leading to submucosal fibrosis (Figure 1). In CD-associated strictures, the areas of submucosal fibrosis are pronounced and accompanied by hyperplasia of the muscularis mucosa,⁵⁰ which can thicken to such a degree that complete obliteration of the submucosa occurs, referred to as obliterative muscularization of the submucosa (OMUS; Figure 1). In addition to hyperplasia, the muscularis mucosa exhibits fiber disarray and ECM deposition. Collectively these changes can represent up to 50% of the increased wall thickness in a CD small bowel stricture.⁵¹ Within the submucosa, hypertrophic nerves are often present, and arteries and veins can show fibromuscular hyperplasia⁵¹ (Figure 1). The muscularis propria is also thickened and expanded by collagen septae.⁵¹⁻⁵³ Fat wrapping, or creeping fat, a major pathologic feature frequently associated with small bowel strictures, is defined by the presence of mesenteric fat extending along the anti-mesenteric border of the intestine.⁵⁴

Scoring Systems for Inflammation in Stricturing Small Bowel Crohn's Disease

Substantial heterogeneity was observed in the grading schemes for inflammation used by the studies included in this review. Most studies graded histologic inflammation using simple ordinal three- or four-grade-scales (Table 1, Supplemental Table 2), which included items that measure active inflammation as well as chronicity.

All scoring systems assessed inflammatory items on H&E stained sections; this was either explicitly stated or inferred.^{14, 27, 32} The terminology used for inflammation varied widely among scoring systems (Supplemental Table 3). Three studies did not specify neutrophilic or chronic inflammation and used the general term "inflammation".^{22, 27, 35} Dillman, *et al.* used the general terms "leukocyte" and "inflammation" and also refers to "mononuclear cells", but did not specify neutrophils.³¹

In the scores that assessed more detailed items in the category of active or neutrophilic inflammation, most systems included ulcers and fissures.^{14, 30, 33, 34, 36, 37} Erosion was part of 3 systems,^{33, 34, 37} and aphthous ulcer was included in 2 systems.^{14, 37} Wagner, *et al.* and Jacene, *et al.* only included "neutrophils".^{17, 38}

Chronic inflammation items also encompassed various terminologies, including chronic inflammatory cells/infiltrates,^{17, 33, 34} lymphocytes,³² lymphoplasmacytic cells,³⁸ and mononuclear cells.^{30, 31} Four systems included "lymphoid aggregates" as an item.^{30, 33, 34, 36} Transmural inflammation was part of most of the scores.^{14, 17, 22, 31, 33, 34, 37}

Regarding other inflammatory cell types, Girlich, *et al.* and Chen, *et al.* included eosinophils.^{30, 32} Granulomas are included in 5 scoring systems.^{32–34, 36, 37}

Within the inflammation component of the scoring system, about a third of the studies *only* included items specifically pertaining to inflammatory *cells* (Supplemental Table 2).^{14, 22, 37, 38} Other items included that pertain to neither inflammatory cells nor fibrosis were edema,^{22, 32, 34, 36} pyloric gland metaplasia,³⁶ villous atrophy or shortening,³⁶ mucin depletion or loss of goblet cells,^{31, 33} crypt architectural changes,^{30, 32} surface epithelial damage,³⁵ increased vascular density,³¹ and pseudopolyps, regenerative epithelium, crypt atrophy, and hemorrhage (Supplemental Table 4).³²

Pucilowska, *et al.*, Jacene, *et al.*^{35, 38} and Pellino, *et al.*,²⁷ used inflammatory scales that only included the *degree* of inflammation, for example mild, moderate, severe (Supplemental Table 2). Adler, *et al.* and Wagner, *et al.* used inflammatory scales that only included the *depth* of inflammation, for example mucosa, submucosa, etc.^{17, 22} The majority of scoring systems used inflammatory scales that include both the degree and the depth of inflammation^{14, 32–34, 36, 37} while 2 investigators did not specify.^{30, 31} Among these, 3 score depth separately from degree,^{30, 32, 36} 3 include depth with degree as one score,^{14, 31, 37} and 2 indicate assessment of depth only through use of the terms transmural inflammation and fissuring ulcers.^{33, 34}

In summary, existing scoring systems show a great heterogeneity regarding the assessment of inflammation. The majority of proposed scoring systems differentiate active from chronic inflammatory alterations employing a graded scoring system; however, the evaluated characteristics including cellular infiltrates as well as various additional criteria differ widely among published studies. Furthermore, while some studies only include the degree of inflammation, but neglect to determine the depth of inflammation, other studies even did not mention the use of any detailed criteria for the assessment of inflammation. Taking the latter into account, none of the published indices for scoring inflammation in surgical resection specimens has been created according to generally accepted principles for the development of evaluative indices and all lack critical validation constituents. A validated index should combine validity (measurement must assess the outcome that it is intended to determine); responsiveness (able to identify substantial alterations according to disease status) and reliability (stable results for repetitive measurements in patients with unchanged disease status).⁵⁵ Ideally, the evaluative tool is easy to use in clinical trials and could reduce the number of patients needed through enhanced discriminatory capacity.⁵⁶ Additionally, the ideal system would incorporate fundamental concepts of disease progression into item scoring, for example cryptitis progresses to crypt abscess, then erosion, then ulceration, and would avoid placing unnecessary limitations on unknown factors, for example that transmural lymphoid aggregates must be present later because they are deep in the bowel wall. Other factors that must be addressed include determining whether quasi-numeric assessments of the amount of microscopic inflammation per low or high power field, or per a certain number of crypts has clinical relevance. Additionally, owing to the substantial heterogeneity that can occur along the length of a stricture, established methods to match histopathologic grading to imaging findings (including surgical handling and potentially processing of the specimen) should be established.

Scoring Systems for Fibrosis in Stricturing Small Bowel Crohn's Disease

For the assessment of fibrostenotic alterations within small bowel CD-associated strictures, heterogeneous scoring systems are available. None of the scores are fully validated to date (Table 1). The majority of the scoring systems use H&E stained sections to evaluate fibrosis items (Supplemental Table 5).^{14, 22, 27, 30, 32–34, 36, 37} Masson's trichrome (MT) was used in systems^{31, 35, 38} and both smooth muscle actin (SMA) staining and Sirius red (SR) was used by one.¹⁷ Chen *et al.* used MT and SMA to grade fibrosis and smooth muscle components, respectively, and compared to grades obtained on H&E stained sections.³⁰ Of note, many studies did not specifically assess stricture and adjacent bowel from the same specimen (Supplemental Table 1).^{17, 22, 31, 33, 38} This is consistent with the concept that the histologic features of strictured small bowel CD and non-strictured small bowel CD overlap, and that an ideal scoring system would include items that are found in both settings, in order to aid in comparison of strictured segments to non-strictured segments.

The fibrosis scoring systems of the included studies encompassed a variety of descriptions for location of fibrosis within the intestinal wall, how it was assessed, and where fibrosis items presented along degree scales. Most of the fibrosis items were assessed using dichotomic and ordinal 3 or 4 grade scales (Supplemental Table 5). One system used an ordinal 5 grade scale.²² Two used other methods of evaluation. Specifically, Jacene, *et al.* assessed the percentage of fibrosis in the area most affected by fibrosis, based on density, intensity, and extent of involvement.³⁸ Wagner *et al.* applied SR staining and immunohistochemical analysis of SMA expression to assess muscular hypertrophy.¹⁷ Using these two staining techniques, they performed software-based calculations of the affected intramural areas of the small bowel CD stricture and determined a ratio between SMA and SR, with a ratio below one indicating increased fibrosis, and a ratio above one reflecting increased muscular hypertrophy.¹⁷

The most commonly included fibrosis item was submucosal fibrosis, which was included in 9 scoring systems (Supplemental Table 6).^{14, 22, 30, 31, 33, 34, 36, 37} Mucosal fibrosis was included in 2 scoring systems^{30, 31} and subserosal fibrosis was included in 3 scoring systems.^{22, 30, 36} Transmural fibrosis or scar was included as an item in 4 scoring systems.^{14, 22, 31, 37} Six studies included an assessment of fibrosis without specifying which bowel wall layer was being assessed.^{17, 27, 30, 32, 35} Regarding the muscularis mucosa and muscularis propria, 3 scoring systems include muscularis mucosa fibrosis as an item^{33, 34, 57} and 4 scoring systems include muscularis propria fibrosis as an item.^{22, 30, 36, 38} Three scoring systems, included the items submucosal fibrosis, muscularis propria fibrosis, and subserosal fibrosis,^{22, 30, 36} and three other scoring systems included the items muscularis mucosa fibrosis and submucosal fibrosis as items.^{33, 34}

Seven scoring systems included items related to these smooth muscle layers that were not specifically "fibrosis" (Supplemental Table 6).^{14, 17, 22, 30, 35, 36, 38} In detail, Smedh, *et al.* included mm thickening,³⁶ Pucilowska, *et al.*, Jacene, *et al.* and Adler, *et al.* included MP thickening,^{22, 35, 38} Chiorean, *et al.* included muscle hyperplasia, but did not specify which muscle layer was assessed,¹⁴ Wagner, *et al.* included muscle hypertrophy, but did not specify which muscle layer was assessed,¹⁷ and Chen, *et al.* included muscle hypertrophy of the

muscularis propria and muscle hyperplasia of the mucosa and submucosa, which is presumed to be referring to the muscularis mucosa.³⁰

The majority of scoring systems only included fibrosis and muscular items and did not consider other ECM alterations.^{14, 22, 27, 31–34, 37} Two scoring systems included neuronal hyperplasia as an item^{30, 33} and one also included adipocyte hyperplasia (Supplemental Table 6).³⁰

Taken together, fibrosis scoring has evolved during the last decade and current histological scoring systems assessing the degree of fibrosis in small bowel CD-associated strictures offer the opportunity to at least roughly estimate the degree of fibrosis. However, it has to be taken into account that none of the above-mentioned scoring systems has been validated, and although the analysis of collagen deposition was one of the main histological items recently evaluated, many studies considered additional histological features, which were not consistent among the studies (Table 1). Reliability of the items describing fibrosis has not been determined, which is particularly important when using grades that are poorly defined, such as ‘mild’, ‘moderate’ or ‘severe’. In the future, an advanced histological scoring system could assess several histological features such as: the quantity and extent of collagen deposition in each layer of the bowel wall (possibly using special stains like MT or SR, similar to the scoring system described by Chen *et al.*³⁰), determination of the degree of preservation or effacement of individual bowel wall layers, the evaluation of fibrotic septae and scar development (for example, refining the concept from the fibrosis scoring system described by Chiorean *et al.*¹⁴), and especially the determination of fibroblast or muscularis propria proliferation/hyperplasia.

DISCUSSION

This systematic review demonstrates considerable heterogeneity in the histologic scoring systems that have been used to assess the severity of the inflammatory and fibrosing components of full thickness small intestine in CD. Importantly, none of the existing indices have been constructed according to accepted methodological standards for the development of evaluative indices and basic knowledge regarding their operating properties is lacking. Furthermore, while some scores have a triple component structure (inflammation, fibrosis and muscle hypertrophy/hyperplasia)^{22, 38} or include the study of the muscle in every bowel wall layer³⁰, specific indices evaluating the important pathological component of myofibroblast hypertrophy/hyperplasia have not been proposed. This circumstance is challenging given that histopathologic evaluation of inflammation, fibrosis, and myofibroblast numbers on surgical resection specimens have and will be used to develop and validate the non-invasive imaging modalities that are likely to be the basis for assessment of outcomes in clinical trials in these patients. Accordingly, development and validation of a comprehensive evaluative index to quantify the process of stenosis histologically is a rate limiting step to future progress in this area and should be considered a top research priority.

It should be acknowledged, however, that research into the histopathology of fibromuscular stenotic disease faces several important challenges. First, a major limitation to performance of human studies is the transmural nature of the pathological process. Therefore, serial

endoscopic biopsies of the mucosa will likely provide little to no information regarding disease progression in deeper mural layers. A second challenge is that the pathological process in stricturing CD is complex and is comprised of multiple elements that include inflammation, fibrosis consisting of pathological deposition of ECM and myofibroblast hypertrophy/hyperplasia. An ideal index should incorporate items that evaluate all of these processes separately as sub-indices and collectively as a global score so that the effects of an intervention can be specifically assessed on the component processes and the overall lesion. Examples on how to evaluate the thickening of the muscularis mucosa or propria could include caliper measurements of thickness, volumetric assessments of the muscle layer, inclusion of specialty stains, such as SMA or MT. Although the concepts of muscle hyperplasia and muscle hypertrophy are well understood, there is currently no consensus opinion on how to quantify either entity in the muscularis mucosa or muscularis propria. In the setting of CD, hyperplasia of the muscularis mucosa may be the only item that can be assessed on H&E (or MT or SMA stain) due to the clearly recognizable deviation from normal, at least in its moderate and extreme state. Automated quantification of muscle cells in the respective layers may be the only option, although it is technically cumbersome.^{58–60} Given the fact that the distribution of muscularis propria thickness measurements in health and disease have not been established, an ideal score could include comparison of the strictured segment with a healthy adjacent bowel segment.

It is noteworthy that most of the studies we reviewed used hematoxylin & eosin (H&E) stained sections to evaluate the inflammatory, fibrotic and muscular components of lesions. Although H&E stained sections are optimal for evaluating inflammation, it may not be an optimal approach for evaluating fibrotic and muscular components in the research setting, given that more sensitive and specific methods are available. Of these methods, SR and MT are the most widely used,^{17, 31, 35, 38} with MT being preferable to SR given that it distinguishes smooth muscle from collagen. Interestingly, in a recent study evaluating fibrotic tissue changes in UC, no difference was noted for the assessment of submucosal fibrosis when using H&E, MT or SR and all stains were highly correlated.⁶¹ Continuous measurement of outcome histopathologic variables (as opposed to semiquantitative or dichotomous scoring systems) would be of great benefit, as many of the imaging variables that are emerging such as magnetization transfer ratio and ultrasound elastography provide continuous variables.

In addition to the issue of heterogeneity in the type of pathological process, stenosis is highly variable in the anatomical distribution of its components. Surprisingly, we found considerable variability in the scoring systems with respect to the mural compartments that were assessed. There is certainly agreement that fibrosis of the submucosa is an important component, and that alterations of the smooth muscle compartment occur. The role of mesenteric fat in structuring CD has been the subject of recent studies,^{54, 62} and interestingly, only one scoring system included adipocyte hyperplasia.³⁰ It is unclear in which sequence bowel wall layers in CD are progressively affected by fibrosis. Hence, a pragmatic approach would be (1) to evaluate all of the mural layers separately, (2) evaluate the degree of correlation for each of the component pathological processes separately (e.g. fibrosis, muscle hyperplasia, adipocyte hyperplasia, etc.) between the mural layers, and (3) subsequently determine whether aggregate or individual scoring of layers is most

appropriate. Ultimately, we believe, an ideal CD stenosis scoring system should account for degree of histopathologic changes specific to each mural layer: hypertrophy/hyperplasia and fibrosis in the muscularis mucosa, fibrosis and muscularization in the submucosa, hypertrophy/hyperplasia and fibrosis in the muscularis propria, and fibrosis in the subserosa. A four-tiered scoring system (e.g. none, mild, moderate, severe) applied to each site (i.e. mucosa, muscularis mucosa, submucosa, muscularis propria, and subserosa) could be ideal.

Finally, as noted previously, none of the scoring systems that we describe have undergone appropriate validation studies. Index validation in stenosing CD has unique challenges that are inherent to the aforementioned sampling constraints. As an initial step reliability testing should be performed to determine inter and intra-observer agreement by pathologists for candidate items. To perform this study standardized item definitions generated by expert pathologists according to a UCLA/RAND process will be required.⁶³ Subsequently construct validation and responsiveness studies are required to define and validate a prototypic evaluative index. A particularly challenging aspect of this paradigm in the context of stenosing CD is responsiveness testing. Responsiveness to change, which is a critical operating property of an evaluative index, is measured by calculating the standardized effect size, which is the degree of change observed following an intervention divided by the amount of measurement variance. Ideally, the data needed to calculate a standardized effect size should come from a placebo controlled trial of a treatment of known efficacy. However, in the case of stenosing CD, no pharmacological agents have been shown to be effective and, at the end of such trials, full thickness resection specimens are likely not available for all patients. In this circumstance patients can be classified as changed and unchanged over time using an external benchmark with face validity such as mural thickening on MRI. Comparable heterogeneity in scoring systems of stricturing small bowel CD exists on cross sectional imaging³. The Stenosis Therapy and Research (STAR) Consortium, consisting of a global group of experts, industry, and regulators, is well under way standardizing and validating novel scoring systems for MRI using the above-mentioned methodology. Ultimately both the histopathology index and the radiology index will then be tested together to serve as a benchmark and allow their use in clinical trials.

Appropriate recommendations for scaling will come from statistical analysis of data from a responsiveness study. Given that CD has a patchy distribution, sampling schemes should be included in index development to define the optimal number of cross sections per mural segment and to determine whether sampling should be different for grossly strictured areas and adjacent tissues. In addition, the concept of a “unified score” in distinction to multiple sub-scores that assess different mural components and different pathological processes must be addressed.

Placebo-controlled trials of anti-fibrotic therapy will require optimal treatment of inflammation in both experimental groups. Therefore, it is critical to assess the inflammatory component of a CD stricture in addition to the components of ECM, collagen and myofibroblast hypertrophy/hyperplasia.² The histopathologic type and severity of inflammation in a stricture may be distinct from inflammation in non-strictured areas and hence warrants separate consideration. Our review of inflammation scoring systems in full-thickness CD revealed heterogeneity. This is best highlighted by grouping the scoring

systems by assessment of the degree and depth of inflammation. The use of diverse terminology and the inclusion of items other than specifically inflammatory cells are also interesting aspects to consider when designing an ideal inflammation scoring system. There is a general pathophysiologic basis for including a scaled progression of neutrophilic inflammation starting with the lamina propria, to cryptitis (intra-epithelial neutrophils), crypt abscess, erosion, ulcer, and fissure, although each of these items may or may not be present simultaneously in the same histologic section. Assessment of chronic inflammation needs to be distinguished from assessment of chronic mucosal injury, with inflammation referring to infiltrates of inflammatory cells, and injury referring to epithelial damage such as crypt architectural distortion, metaplastic changes, and possibly even lamina propria fibrosis. Also, the role of other inflammatory cells, such as eosinophils, needs further study. We did not include in our review other CD inflammation scoring systems performed in the setting of biopsy or in the absence of concomitant fibrosis scoring, as, ultimately, radiologic correlation will depend on both inflammatory and fibromuscular stenosis scoring in full thickness resections.

CONCLUSION

In summary, stenosing CD is an important medical problem that is inadequately managed by existing medical therapy. Progress in this field will require the development of specific treatments to address the related yet distinct entities of inflammation, fibrosis and muscular hypertrophy. Although multiple histologic scoring systems have been described to quantify fibrosis and inflammation in small bowel CD strictures, they have not been created according to established methods for evaluative index development, and their operating properties are unknown. Current histologic scoring systems reveal heterogeneous definitions of fibrosis and are incongruent as to the importance of fibrosis and mesenchymal cell hyperplasia in each of the bowel wall layers. Furthermore, very little emphasis has been placed on the evaluation of muscular hypertrophy as a critical pathological component. Development of a fully validated histologic index is essential for validation of novel cross-sectional imaging techniques that may ultimately serve as outcome measures in clinical trials. The importance of developing endpoints for stenosing CD has been recognized by two international working groups^{29, 64}. Proposed endpoints²⁹, including a patient reported outcome tool, an MRI index, and a histopathology score, are being developed under the umbrella of the STAR consortium with involvement of interdisciplinary global expert panels, industry and regulators. A clinical trial protocol has already been provided²⁹. This will allow testing of anti-fibrotic drugs in small bowel stricturing CD, once these indices are available.

An ideal histopathology index should evaluate several elements of stricturing small bowel CD separately: inflammation, fibrosis (defined as the excessive accumulation of ECM) and myofibroblast hypertrophy/hyperplasia. These elements should be determined independently for each layer of the intestinal wall in the form of subscores. The above elements have not been thoroughly defined or quantified and their variability has not been determined in the healthy gut. The same holds true for strictured segments of the CD small intestine. It may hence be most accurate to include a same patient control of a healthy adjacent bowel segment and compare it to the affected segment measuring relative differences. While it

would be ideal to propose a non-validated all-encompassing scoring system in this review article we respectfully refrained from doing so. We feel that a standardized process with item definitions according to a UCLA/RAND process will be required first, followed by reliability testing, both supported by expert pathologists. It may be desirable to have a scoring system for each part of the small bowel separately. We will, however, focus on terminal ileal strictures as this is the likely first population to enter a clinical trial for an anti-fibrotic drug based on an international expert consensus²⁹. These projects have already begun and will be reported separately.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant support: This work was supported by the Helmsley Charitable Trust through the Stenosis Therapy and Anti-Fibrotic Research (STAR) Consortium.

Conflict of interest:

DB is on the advisory board or consultant for Amgen, AbbVie, Dr. Falk Foundation, Ferring, MSD, Pfizer, Pharmacosmos, Roche, Takeda, Tillotts Pharma and Vifor.

CEP and **TN** are employees of Robarts Clinical Trials, Inc.

VJ receives salary support from the John and Susan McDonald Endowed IBD Chair at Western University, London, Ontario, Canada; consulting fees from AbbVie, Eli Lilly, GlaxoSmithKline, Arena pharmaceuticals, Genentech, Pendopharm, Sandoz, Merck, Takeda, Janssen, Robarts Clinical Trials, Topivert, Celltrion; speaker's fees from Takeda, Janssen, Shire, Ferring, Abbvie, Pfizer

JGF receives grants to his institution from Siemens Healthineers and Medtronic.

BGF is has received grant/research support from Millennium Pharmaceuticals, Merck, Tillotts Pharma AG, AbbVie, Novartis Pharmaceuticals, Centocor Inc., Elan/Biogen, UCB Pharma, Bristol-Myers Squibb, Genentech, ActoGenix, and Wyeth Pharmaceuticals Inc.; consulting fees from Millennium Pharmaceuticals, Merck, Centocor Inc., Elan/Biogen, Janssen-Ortho, Teva Pharmaceuticals, Bristol-Myers Squibb, Celgene, UCB Pharma, AbbVie, Astra Zeneca, Serono, Genentech, Tillotts Pharma AG, Unity Pharmaceuticals, Albireo Pharma, Given Imaging Inc., Salix Pharmaceuticals, Novonordisk, GSK, Actogenix, Prometheus Therapeutics and Diagnostics, Athersys, Axcan, Gilead, Pfizer, Shire, Wyeth, Zealand Pharma, Zyngenia, GiCare Pharma Inc., and Sigmoid Pharma; and speakers bureau fees from UCB, AbbVie, and J&J/Janssen

FR is on the advisory board or consultant for AbbVie, Allergan, Celgene, Gossamer, Receptos, Thetis, UCB, Samsung, Koutif, Pliant, Boehringer-Ingelheim, Metacrine, Takeda, Allergan, Helmsley, RedX, Gilead and Roche.

SK is supported by National Institute of Health grant 5T32DK083251-08

PB reports personal fees from MSD, personal fees from Roche, and personal fees from Astra-Zeneca, outside the submitted work

MR reports Merck- speaker panel; Bayer- Pathologist on a clinical trial; Chief Scientific Officer of Beyond Celiac, a non-profit patient support organization, outside of the submitted work.

GDH has received consulting fees from Glaxo Smith Kline, Shire Pharmaceuticals, Teva Pharma, Galapagos, Genentech, Novartis Pharma, Fast Forward Pharmaceuticals, Takeda and Janssen R&D, and University Hospital Leuven has received support for her from Centocor and Takeda.

RKP has received consulting fees from Seres Therapeutics, Genentech, Eli Lilly, Protagonist, Arena, and Gossamer Bio.

NH is a consultant for Abbvie, Celgene and Lilly USA.

DHB is a consultant for and receives research support from Medtronic.

MEB receives no direct support. Cleveland Clinic receives support for him from Siemens Healthineers in the form of salary, software and hardware for the investigation of reduced exposure in CT Enterography

IOG receives no direct support. Cleveland Clinic receives support for her from UCB, Celgene, and Pliant Therapeutics

No conflict of interest: RM, AB, CR, RKP, MAV, RF, RO

REFERENCES

1. Rieder F, Fiocchi C, Rogler G. Mechanisms, Management, and Treatment of Fibrosis in Patients With Inflammatory Bowel Diseases. *Gastroenterology* 2017;152:340–350 e6. [PubMed: 27720839]
2. Rieder F, Latella G, Magro F, et al. European Crohn's and Colitis Organisation Topical Review on Prediction, Diagnosis and Management of Fibrostenosing Crohn's Disease. *J Crohns Colitis* 2016;10:873–85. [PubMed: 26928961]
3. Bettenworth D, Bokemeyer A, Baker M, et al. Assessment of Crohn's disease-associated small bowel strictures and fibrosis on cross-sectional imaging: a systematic review. *Gut* 2019;68:1115–1126. [PubMed: 30944110]
4. Bettenworth D, Gustavsson A, Atreja A, et al. A Pooled Analysis of Efficacy, Safety, and Long-term Outcome of Endoscopic Balloon Dilation Therapy for Patients with Stricture Crohn's Disease. *Inflamm Bowel Dis* 2017;23:133–142. [PubMed: 28002130]
5. Bettenworth D, Mucke MM, Lopez R, et al. Efficacy of Endoscopic Dilation of gastroduodenal Crohn's disease strictures: A Systematic Review and Meta-analysis of Individual Patient Data. *Clin Gastroenterol Hepatol* 2018.
6. Bemelman WA, Allez M. The surgical intervention: earlier or never? *Best Pract Res Clin Gastroenterol* 2014;28:497–503. [PubMed: 24913388]
7. Bettenworth D, Rieder F. Medical therapy of stricturing Crohn's disease: what the gut can learn from other organs - a systematic review. *Fibrogenesis Tissue Repair* 2014;7:5. [PubMed: 24678903]
8. Noble PW, Albera C, Bradford WZ, et al. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. *Lancet* 2011;377:1760–9. [PubMed: 21571362]
9. Ferguson MW, Duncan J, Bond J, et al. Prophylactic administration of avotermin for improvement of skin scarring: three double-blind, placebo-controlled, phase I/II studies. *Lancet* 2009;373:1264–74. [PubMed: 19362676]
10. el-Agroudy AE, Hassan NA, Foda MA, et al. Effect of angiotensin II receptor blocker on plasma levels of TGF-beta 1 and interstitial fibrosis in hypertensive kidney transplant patients. *Am J Nephrol* 2003;23:300–6. [PubMed: 12904684]
11. Diez J, Querejeta R, Lopez B, et al. Losartan-dependent regression of myocardial fibrosis is associated with reduction of left ventricular chamber stiffness in hypertensive patients. *Circulation* 2002;105:2512–7. [PubMed: 12034658]
12. Deepak P, Fletcher JG, Fidler JL, et al. Computed Tomography and Magnetic Resonance Enterography in Crohn's Disease: Assessment of Radiologic Criteria and Endpoints for Clinical Practice and Trials. *Inflamm Bowel Dis* 2016;22:2280–8. [PubMed: 27508513]
13. Wilkens R, Hagemann-Madsen RH, Peters DA, et al. Validity of Contrast-enhanced Ultrasonography and Dynamic Contrast-enhanced MR Enterography in the Assessment of Transmural Activity and Fibrosis in Crohn's Disease. *J Crohns Colitis* 2018;12:48–56. [PubMed: 28981627]
14. Chiorean MV, Sandrasegaran K, Saxena R, et al. Correlation of CT enteroclysis with surgical pathology in Crohn's disease. *Am J Gastroenterol* 2007;102:2541–50. [PubMed: 17900329]
15. Jensen MD, Nathan T, Rafaelsen SR, et al. Diagnostic accuracy of capsule endoscopy for small bowel Crohn's disease is superior to that of MR enterography or CT enterography. *Clin Gastroenterol Hepatol* 2011;9:124–9. [PubMed: 21056692]

16. Soyer P, Boudiaf M, Sirol M, et al. Suspected anastomotic recurrence of Crohn disease after ileocolic resection: evaluation with CT enteroclysis. *Radiology* 2010;254:755–64. [PubMed: 20177090]
17. Wagner M, Ko HM, Chatterji M, et al. Magnetic resonance imaging predicts histopathologic composition of ileal Crohn's disease. *J Crohns Colitis* 2018.
18. Punwani S, Rodriguez-Justo M, Bainbridge A, et al. Mural inflammation in Crohn disease: location-matched histologic validation of MR imaging features. *Radiology* 2009;252:712–20. [PubMed: 19635832]
19. Steward MJ, Punwani S, Proctor I, et al. Non-perforating small bowel Crohn's disease assessed by MRI enterography: derivation and histopathological validation of an MR-based activity index. *Eur J Radiol* 2012;81:2080–8. [PubMed: 21924572]
20. Catalano OA, Gee MS, Nicolai E, et al. Evaluation of Quantitative PET/MR Enterography Biomarkers for Discrimination of Inflammatory Strictures from Fibrotic Strictures in Crohn Disease. *Radiology* 2016;278:792–800. [PubMed: 26436860]
21. Torres J, Mehndru S, Colombel JF, et al. Crohn's disease. *Lancet* 2017;389:1741–1755. [PubMed: 27914655]
22. Adler J, Punglia DR, Dillman JR, et al. Computed tomography enterography findings correlate with tissue inflammation, not fibrosis in resected small bowel Crohn's disease. *Inflamm Bowel Dis* 2012;18:849–56. [PubMed: 21710535]
23. Rimola J, Planell N, Rodriguez S, et al. Characterization of inflammation and fibrosis in Crohn's disease lesions by magnetic resonance imaging. *Am J Gastroenterol* 2015;110:432–40. [PubMed: 25623654]
24. Ripolles T, Rausell N, Paredes JM, et al. Effectiveness of contrast-enhanced ultrasound for characterisation of intestinal inflammation in Crohn's disease: a comparison with surgical histopathology analysis. *J Crohns Colitis* 2013;7:120–8. [PubMed: 22483566]
25. Tielbeek JA, Ziech ML, Li Z, et al. Evaluation of conventional, dynamic contrast enhanced and diffusion weighted MRI for quantitative Crohn's disease assessment with histopathology of surgical specimens. *Eur Radiol* 2014;24:619–29. [PubMed: 24037299]
26. Baumgart DC, Muller HP, Grittner U, et al. US-based Real-time Elastography for the Detection of Fibrotic Gut Tissue in Patients with Stricture Crohn Disease. *Radiology* 2015;275:889–99. [PubMed: 25668520]
27. Pellino G, Nicolai E, Catalano OA, et al. PET/MR Versus PET/CT Imaging: Impact on the Clinical Management of Small-Bowel Crohn's Disease. *J Crohns Colitis* 2016;10:277–85. [PubMed: 26574490]
28. Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol* 1992;45:1341–5. [PubMed: 1460470]
29. Rieder F, Bettenworth D, Ma C, et al. An expert consensus to standardise definitions, diagnosis and treatment targets for anti-fibrotic stricture therapies in Crohn's disease. *Aliment Pharmacol Ther* 2018;48:347–357. [PubMed: 29920726]
30. Chen W, Lu C, Hirota C, et al. Smooth Muscle Hyperplasia/Hypertrophy is the Most Prominent Histological Change in Crohn's Fibrostenosing Bowel Strictures: A Semiquantitative Analysis by Using a Novel Histological Grading Scheme. *J Crohns Colitis* 2017;11:92–104. [PubMed: 27364949]
31. Dillman JR, Stidham RW, Higgins PD, et al. Ultrasound shear wave elastography helps discriminate low-grade from high-grade bowel wall fibrosis in ex vivo human intestinal specimens. *J Ultrasound Med* 2014;33:2115–23. [PubMed: 25425367]
32. Girlich C, Jung EM, Huber E, et al. Comparison between preoperative quantitative assessment of bowel wall vascularization by contrast-enhanced ultrasound and operative macroscopic findings and results of histopathological scoring in Crohn's disease. *Ultraschall Med* 2011;32:154–9. [PubMed: 20449794]
33. Kotanagi H, Kramer K, Fazio VW, et al. Do microscopic abnormalities at resection margins correlate with increased anastomotic recurrence in Crohn's disease? Retrospective analysis of 100 cases. *Dis Colon Rectum* 1991;34:909–16. [PubMed: 1914726]

34. Maconi G, Carsana L, Fociani P, et al. Small bowel stenosis in Crohn's disease: clinical, biochemical and ultrasonographic evaluation of histological features. *Aliment Pharmacol Ther* 2003;18:749–56. [PubMed: 14510749]
35. Pucilowska JB, McNaughton KK, Mohapatra NK, et al. IGF-I and procollagen alpha1(I) are coexpressed in a subset of mesenchymal cells in active Crohn's disease. *Am J Physiol Gastrointest Liver Physiol* 2000;279:G1307–22. [PubMed: 11093955]
36. Smedh K, Olaison G, Franzen L, et al. Endoscopic and external bowel changes and histopathology in patients with Crohn's disease. *Br J Surg* 1995;82:191–4. [PubMed: 7749686]
37. Zappa M, Stefanescu C, Cazals-Hatem D, et al. Which magnetic resonance imaging findings accurately evaluate inflammation in small bowel Crohn's disease? A retrospective comparison with surgical pathologic analysis. *Inflamm Bowel Dis* 2011;17:984–93. [PubMed: 20722053]
38. Jacene HA, Ginsburg P, Kwon J, et al. Prediction of the need for surgical intervention in obstructive Crohn's disease by 18F-FDG PET/CT. *J Nucl Med* 2009;50:1751–9. [PubMed: 19837758]
39. Yantiss RK, Odze RD. Diagnostic difficulties in inflammatory bowel disease pathology. *Histopathology* 2006;48:116–32. [PubMed: 16405661]
40. Bressenot A, Geboes K, Vignaud JM, et al. Microscopic features for initial diagnosis and disease activity evaluation in inflammatory bowel disease. *Inflamm Bowel Dis* 2013;19:1745–52. [PubMed: 23782793]
41. Odze R Diagnostic problems and advances in inflammatory bowel disease. *Mod Pathol* 2003;16:347–58. [PubMed: 12692200]
42. Stappenbeck TS, McGovern DPB. Paneth Cell Alterations in the Development and Phenotype of Crohn's Disease. *Gastroenterology* 2017;152:322–326. [PubMed: 27729212]
43. Masterson JC, Capocelli KE, Hosford L, et al. Eosinophils and IL-33 Perpetuate Chronic Inflammation and Fibrosis in a Pediatric Population with Stricture Crohn's Ileitis. *Inflamm Bowel Dis* 2015;21:2429–40. [PubMed: 26218140]
44. Cayci M, Bostanci EB, Turhan N, et al. The analysis of clinico-pathologic characteristics in patients who underwent surgery due to stricturing and non-perineal fistulizing forms of Crohn's disease: a retrospective cohort study. *Int J Surg* 2015;15:49–54. [PubMed: 25647540]
45. Rotterdam H, Korelitz BI, Sommers SC. Microgranulomas in grossly normal rectal mucosa in Crohn's disease. *Am J Clin Pathol* 1977;67:550–4. [PubMed: 868794]
46. Bressenot A, Peyrin-Biroulet L. Histologic features predicting postoperative Crohn's disease recurrence. *Inflamm Bowel Dis* 2015;21:468–75. [PubMed: 25437814]
47. Sokol H, Polin V, Lavergne-Slove A, et al. Plexitis as a predictive factor of early postoperative clinical recurrence in Crohn's disease. *Gut* 2009;58:1218–25. [PubMed: 19625280]
48. Lemmens B, de Buck van Overstraeten A, Arijs I, et al. Submucosal Plexitis as a Predictive Factor for Postoperative Endoscopic Recurrence in Patients with Crohn's Disease Undergoing a Resection with Ileocolonic Anastomosis: Results from a Prospective Single-centre Study. *J Crohns Colitis* 2017;11:212–220. [PubMed: 27466173]
49. Ferrante M, de Hertogh G, Hlavaty T, et al. The value of myenteric plexitis to predict early postoperative Crohn's disease recurrence. *Gastroenterology* 2006;130:1595–606. [PubMed: 16697723]
50. Koukoulis G, Ke Y, Henley JD, et al. Obliterative muscularization of the small bowel submucosa in Crohn disease: a possible mechanism of small bowel obstruction. *Arch Pathol Lab Med* 2001;125:1331–4. [PubMed: 11570909]
51. Zhang X, Ko HM, Torres J, et al. Luminally polarized mural and vascular remodeling in ileal strictures of Crohn's disease. *Hum Pathol* 2018;79:42–49. [PubMed: 29555578]
52. Burke JP, Mulrow JJ, O'Keane C, et al. Fibrogenesis in Crohn's disease. *Am J Gastroenterol* 2007;102:439–48. [PubMed: 17156147]
53. Graham MF, Diegelmann RF, Elson CO, et al. Collagen content and types in the intestinal strictures of Crohn's disease. *Gastroenterology* 1988;94:257–65. [PubMed: 3335305]
54. Mao R, Kurada S, Gordon IO, et al. The Mesenteric Fat and Intestinal Muscle Interface: Creeping Fat Influencing Stricture Formation in Crohn's Disease. *Inflamm Bowel Dis* 2018.

55. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis* 1985;38:27–36. [PubMed: 3972947]
56. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40:171–8. [PubMed: 3818871]
57. Bataille F, Klebl F, Rummele P, et al. Histopathological parameters as predictors for the course of Crohn's disease. *Virchows Arch* 2003;443:501–7. [PubMed: 12898246]
58. Woodruff PG, Dolganov GM, Ferrando RE, et al. Hyperplasia of smooth muscle in mild to moderate asthma without changes in cell size or gene expression. *Am J Respir Crit Care Med* 2004;169:1001–6. [PubMed: 14726423]
59. Owens GK, Rabinovitch PS, Schwartz SM. Smooth muscle cell hypertrophy versus hyperplasia in hypertension. *Proc Natl Acad Sci U S A* 1981;78:7759–63. [PubMed: 6950415]
60. Blennerhassett MG, Vignjevic P, Vermillion DL, et al. Inflammation causes hyperplasia and hypertrophy in smooth muscle of rat small intestine. *Am J Physiol* 1992;262:G1041–6. [PubMed: 1616034]
61. Gordon IO, Agrawal N, Willis E, et al. Fibrosis in ulcerative colitis is directly linked to severity and chronicity of mucosal inflammation. *Aliment Pharmacol Ther* 2018;47:922–939. [PubMed: 29411405]
62. Coffey CJ, Kiernan MG, Sahebally SM, et al. Inclusion of the Mesentery in Ileocolic Resection for Crohn's Disease is Associated With Reduced Surgical Recurrence. *J Crohns Colitis* 2018;12:1139–1150. [PubMed: 29309546]
63. Fitch K, Bernstein SJ, Aguilar MD, et al. The RAND/UCLA Appropriateness Method User's Manual: RAND, 2001.
64. Danese S, Bonovas S, Lopez A, et al. Identification of Endpoints for Development of Antifibrosis Drugs for Treatment of Crohn's Disease. *Gastroenterology* 2018;155:76–87. [PubMed: 29601825]

What you need to know:**BACKGROUND AND CONTEXT:**

We performed a systematic review of all published histopathologic scoring systems used to assess stenosing Crohn's disease.

NEW FINDINGS:

We found a lack of validated histopathologic indices for assessment of fibromuscular stenosis.

LIMITATIONS:

This was a review of previously published research.

IMPACT:

Data that describe the operating properties of cross-sectional imaging techniques for stenosing CD and use the histopathology as the reference standard should be questioned. Development and validation of a scoring system based on imaging findings is an important research priority.

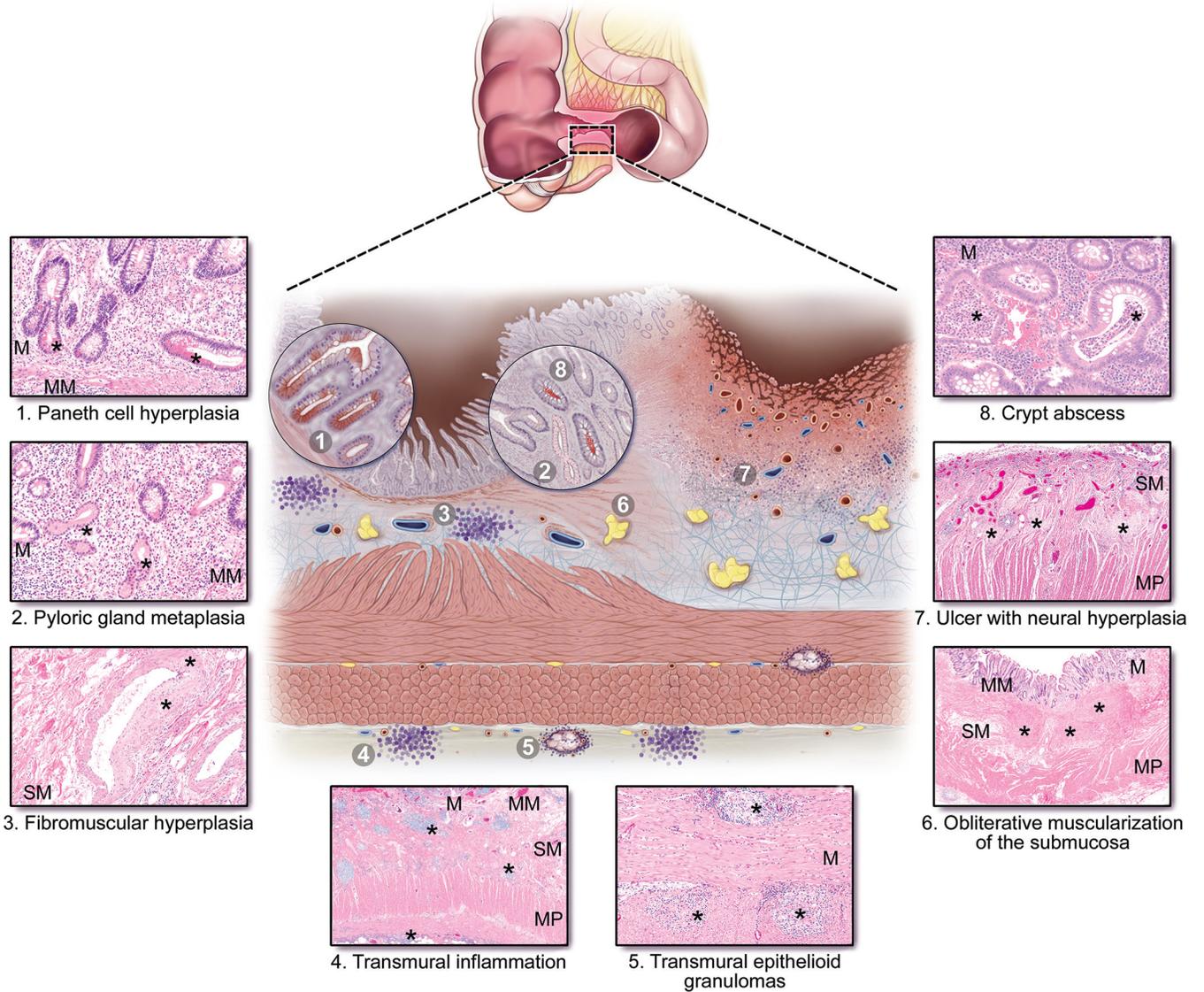


Figure 1: Typical histopathologic features found in Crohn's disease associated strictures. Typical histopathologic findings are described below each histologic image and marked by (*) and are additionally depicted in the composite illustration (center). Original magnification clockwise from top right is: 20x, 4x, 2x, 10x, 2x, 10x, 10x, and 10x. Abbreviations: M, mucosa; MM, muscularis mucosa; SM, submucosa; MP, muscularis propria.

Published pathological scoring systems assessing inflammation and fibrosis in Crohn's disease patients with small bowel disease manifestation.

Table 1:

Author, year ^{Ref}	Study design	No of patients included (n)	Strictureing disease phenotype (%)	Criteria/Scoring system used to define small bowel inflammation	Criteria/Scoring system used to define small bowel fibrosis	Comparing technique (for fibrosis/inflammation)
Kotnagi 1991 ³³	Retrospective cohort	100	n. a.	Ordinal 3 grade scale (absent, mild or focal, diffuse or marked) for 14 items: Mucosal/submucosal edema, neutrophils in mucosa/submucosa, mucosal chronic inflammatory cells, lymphoid aggregates, pyloric gland metaplasia, fibrosis of muscularis mucosae/submucosa, cryptitis/crypt abscess, mucosal erosions/ulcers, fissural ulcers, granuloma, shortening of intestinal villi, depletion of intracellular mucin, neuronal hyperplasia and transmural inflammation.	Mucosal/submucosal edema, lymphoid aggregates, pyloric metaplasia, ulcers, muscularis mucosa thickened, submucosal edema, submucosal/muscularis propria/subserosal fibrosis, submucosal lymphoid aggregates, submucosal dilated lymph vessels, granuloma and fissures	None
Smedh 1995 ³⁶	Retrospective cohort	18	n. a.	Dichotomic scale (absent/present) for 11 items, with global assessment of severity (0–100): Villous atrophy, epithelial/lamina propria/submucosal/muscularis propria/subserosal leukocyte infiltration, pyloric metaplasia, ulcers, muscularis mucosa thickened, submucosal edema, submucosal/muscularis propria/subserosal fibrosis, submucosal lymphoid aggregates, submucosal dilated lymph vessels, granuloma and fissures	Villous atrophy, ulcers, muscularis mucosa thickened, submucosal edema, submucosal/muscularis propria/subserosal fibrosis, submucosal lymphoid aggregates, submucosal dilated lymph vessels, granuloma and fissures	Endoscopy
Pucitowska 2000 ³⁵	Prospective cohort	11	100%	Ordinal 3 grade scale (none, mild, severe) for 4 items, summed to a total histology score: Surface epithelial damage, lamina propria inflammation, thickness of muscularis propria and fibrosis.	Surface epithelial damage, lamina propria inflammation, thickness of muscularis propria and fibrosis.	None
Maconi 2003 ³² , adapted from Kotnagi ³³	Prospective cohort	43	100%	- Ordinal 4 grade scale (none, mild, moderate, severe) for 8 items: Mucosal/submucosal edema, chronic inflammatory infiltrate, neutrophil infiltrate, crypt abscesses, mucosal erosions, wall ulcers, lymphoid aggregates and fibrosis of muscularis mucosae and submucosa. - Dichotomic scale (absent/present) for 3 items: Transmural inflammation, fissural ulcers and non-necrotizing granulomas	Mucosal/submucosal edema, chronic inflammatory infiltrate, neutrophil infiltrate, crypt abscesses, mucosal erosions, wall ulcers, lymphoid aggregates and fibrosis of muscularis mucosae and submucosa. - Dichotomic scale (absent/present) for 3 items: Transmural inflammation, fissural ulcers and non-necrotizing granulomas	Ultrasound, CDAL, CRP
Chiorean 2007 ¹⁴	Prospective cohort	44	66%	Ordinal 4 grade scale: 0) No positive features 1) Mild: Aphthous ulcers affected surface <50%; cryptitis <50%, inflammation limited to mucosa 2) Moderate: Large, superficial ulcers (0.5–2cm); ulcerated surface <50%, affected surface 50–100%; cryptitis >50%; crypt abscess; submucosal inflammation 3) Severe: Deep Ulcers or size >2cm; circumferential ulcers; transmural inflammation; deep fissures. - Lesions are classified as predominantly inflammatory if inflammatory score > 1 and fibrostenosis score 1. - Lesions are classified as predominantly fibrostenotic if fibrostenosis > 1 and inflammatory score 1 - Compound lesions classified if inflammatory and fibrostenosis scores 1	Ordinal 3 grade scale: 0) None: No or minimal fibrosis limited to submucosa (<25% thickness) 1) Mild/moderate: Mild stricture (>15mm) with nondilated lumen; submucosal fibrosis and muscular hyperplasia >25% with preserved layers 2) Severe: Massive transmural fibrosis; effacement of normal layers; severe stricture.	CT
Jacene 2009 ³⁸	Prospective cohort	13	100%	Ordinal 4 grade scale for two items: acute inflammation (degree of neutrophilic infiltration) and chronic inflammation (degree of lymphoplasmacytic infiltration): 1) Absent 2) Mild 3) Moderate 4) Severe	Percentage of fibrosis involving muscularis propria in the most affected area using trichrome staining based on density, intensity, and extent of involvement (expressed in 10% increments). Additionally, assessment of muscle hypertrophy: determination of the diameter of the thickest hypertrophied muscularis propria compared to normal-appearing muscularis propria in the same patient	PET-CT
Zappa 2011 ³⁷	Retrospective cohort	53	83%	Ordinal 3 grade scale: 0) Nonactive CD: minimal neutrophil infiltrate limited to mucosa	Ordinal 3 grade scale: 0) Minimal fibrosis: fibrosis limited to submucosa 1) Moderate fibrosis: massive submucosal fibrosis	MRI

Author, year ^{Ref}	Study design	No of patients included (n)	Stricture disease phenotype (%)	Criteria/Scoring system used to define small bowel inflammation	Criteria/Scoring system used to define small bowel fibrosis	Comparing technique (for fibrosis/inflammation)
Girlich 2011 ³² , modified from Bataille 2003 ³⁷	Prospective cohort	20	40%	1) Moderately active CD: neutrophil infiltrate limited to mucosa/submucosa without muscular involvement 2) Severely active CD: transmural neutrophil infiltrate and/or fistula and/or abscesses in subserosa Ordinal grade scale for 3 items with subgroups: - mucosal surface (neutrophils, eosinophils, lymphocytes, goblet cells, erosion) - Depth of wall infiltration (neutrophils, eosinophils, lymphocytes) - Density of wall infiltration (neutrophils, eosinophils, lymphocytes) Dichotomous scale (present/absent) for another 9 items: - Pseudopolyps, regenerative epithelium, crypt architecture, crypt abscess, crypt atrophy, fibrosis, edema, hemorrhage, epithelioid granulomas	with preserved layers 2) Severe fibrosis: massive transmural fibrosis with effacement of normal layers.	Ultrasound
Adler 2012 ²² adapted from Macconi ³⁴ , Thiess, and Chiorean ¹⁴	Retrospective cohort	22	54.5%	Ordinal 5 grade scale: 0) No inflammation 1) Lamina propria inflammation only 2) Submucosal foci of inflammation 3) Foci of transmural inflammation 4) Significant confluent transmural inflammation.	Ordinal 5 grade scale: 0) No fibrosis 1) Minimal fibrosis in submucosa or subserosa 2) Increased submucosal fibrosis, septa into muscularis propria 3) Septa through muscularis propria, increase in subserosal collagen 4) Significant transmural scar, marked subserosal collagen Dichotomous scale: Presence or absence of increased thickness of muscularis propria	CT
Dillmann 2014 ³¹ adapted from scoring system for experimental murine colitis [*]	Prospective cohort	8	n. a.	Ordinal 5 grade scale: 0) No inflammation 1) Low level of inflammation with scattered infiltrating mononuclear cells 2) Moderate inflammation with multiple foci 3) High level of inflammation with increased vascular density and marked wall thickening 4) Maximal severity of inflammation with transmural leukocyte infiltration and loss of goblet cells.	Ordinal 4 grade scale: 0) No architectural distortion, no abnormal Masson staining 1) No architectural distortion, mild abnormal Masson staining in mucosa/submucosa 2) Substantial abnormal mucosa/submucosal Masson staining with modest distortion of architecture 3) Transmural fibrosis with abnormal Masson staining in all layers, transmural architectural distortion	Ultrasound
Pellino 2016 ²⁷	Prospective cohort	29	At least 50%	Ordinal 3 grade scale: 0) None/mild 1) Moderate 2) Severe	Ordinal 3 grade scale: 0) None/mild 1) Moderate 2) Severe	PET-CT, PET-MRI, MRI
Chen 2017 ³⁰	Retrospective cohort	48	100%	Ordinal 4 grade scale (absent, mild, moderate, severe) for 6 items evaluated for each bowel layer (mucosa, submucosa, muscularis propria and subserosal adventitia): 1) Active inflammation (superficial ulceration, fissuring ulceration, cryptitis, crypt abscess and neutrophil infiltration in lamina propria) 2) Chronic inflammation (lamina propria mononuclear cell cellularity, eosinophil infiltration, lymphoid aggregates and crypt architecture alteration) 3) Fibrosis (extent of fibroblast proliferation mixed with collagen and other matrix deposition forming scar-like fibrous tissue) 4) Smooth muscle hyperplasia (in mucosa and submucosa) and hypertrophy (of muscularis propria) 5) Neuronal hypertrophy 6) Adipocyte hyperplasia		None

Author, year ^{Ref}	Study design	No of patients included (n)	Stricturing disease phenotype (%)	Criteria/Scoring system used to define small bowel inflammation	Criteria/Scoring system used to define small bowel fibrosis	Comparing technique (for fibrosis/inflammation)
Wagner 2018 ¹⁷	Retrospective cohort	35	At least 34%	Ordinal 3 grade scale for depth of neutrophil infiltrates: 1) None or mucosa only 2) Submucosa 3) Muscularis propria/subserosa/Serosa Dichotomic scale for extent of bowel wall edema: 1) No edema or minimal edema 2) Obvious edema	Immunostaining of smooth muscle actin (SMA) for muscular hypertrophy and Sirius red special stain for collagen deposition. Software-based calculation of affected intramural areas. Determination of ratio between SMA and Sirius red was calculated: ratio < 1: increased fibrosis; ratio > 1: increased muscular hypertrophy.	MRI

*: Johnson et al. Inflamm Bowel Dis. 2012;18:460–71 and Higgins et al. Comparative Immunology, Microbiology and Infectious Diseases. 2011;34:247–57.