6-1-2021

An International Consensus to Standardize Integration of Histopathology in Ulcerative Colitis Clinical Trials

Christopher Ma
Cumming School of Medicine

Rocio Sedano
Western University

Ahmed Almradi
Western University

Niels Vande Casteele
Robarts Research Institute

Claire E. Parker
Robarts Research Institute

See next page for additional authors

Authors

Christopher Ma, Rocio Sedano, Ahmed Almradi, Niels Vande Casteele, Claire E. Parker, Leonardo Guizzetti, David F. Schaeffer, Robert H. Riddell, Reetesh K. Pai, Robert Battat, Bruce E. Sands, Christophe Rosty, Marla C. Dubinsky, Florian Rieder, Noam Harpaz, Maria T. Abreu, Robert V. Bryant, Gregory Y. Lauwers, Richard Kirsch, Mark A. Valasek, Eileen Crowley, William J. Sandborn, Brian G. Feagan, Rish K. Pai, and Vipul Jairath

CORRESPONDING AUTHOR: Dr. Christopher Ma, MD, MPH, Assistant Professor, Division of Gastroenterology & Hepatology, Departments of Medicine and Community Health Sciences, Cumming School of Medicine, University of Calgary, 3280 Hospital Drive NW, Calgary, Alberta, Canada, T2N 4Z6, Tel: 403-220-8457, Fax: 403-592-5090, christopher.ma@ucalgary.ca.

AUTHOR CONTRIBUTIONS:

Conception and design: CM, RiKP, WJS, BGF, VJ

Data collection: all authors

Data analysis and interpretation: CEP, LG

Manuscript drafting: CM, CEP, RiKP, BGF, VJ

Manuscript editing for important intellectual content: all authors

Authorship Statement: All authors have reviewed, contributed to, and approved the final version of this manuscript.

Article Guarantor: CM and VJ are serving as the article guarantors.

Conflicts of Interest:

CM has received consulting fees from AbbVie, AVIR Pharma Inc, Janssen, Takeda, Pfizer, Roche, and Alimentiv Inc. (formerly Robarts Clinical Trials Inc.); speaker's fees from AbbVie, Janssen, Takeda, and Pfizer; and research support from the Canadian IBD Research Consortium and Pfizer.

RS has no conflicts of interest to report.

AA has no conflicts of interest to report.

NVC has received research grants from R-Biopharm; grants and personal fees from Takeda and UCB; and personal fees from Alimentiv, Inc. (formerly Robarts Clinical Trials, Inc.), Celltrion and Prometheus.

CEP is an employee of Alimentiv Inc. (formerly Robarts Clinical Trials, Inc.).

LG is an employee of Alimentiv Inc. (formerly Robarts Clinical Trials, Inc.).

DFS has received honoraria from Alimentiv Inc, Pfizer, Amgen, Merck and Diaceutics, and stock ownership from Satisfai Health.

RHR has no conflicts of interest to report.

ReKP has received consulting fees from Alimentiv, Inc. (formerly Robarts Clinical Trials, Inc.).

RB has received grant funding from Weill Cornell Medicine Department of Medicine Fund for the Future Program and receive research funding from the Jill Roberts IBD Research Fund.

BES discloses consulting fees from 4D Pharma, Abbvie, Allergan, Amgen, Arena Pharmaceuticals, AstraZeneca, Boehringer Ingelheim, Boston Pharmaceuticals, Capella Biosciences, Celgene, Celltrion Healthcare, EnGene, Ferring, Genentech, Gilead, Hoffmann-La Roche, Immunic, Ironwood Pharmaceuticals, Janssen, Lilly, Lyndra, MedImmune, Morphic Therapeutic, Oppilan Pharma, OSE Immunotherapeutics, Otsuka, Palatin Technologies, Pfizer, Progenity, Prometheus Laboratories, Redhill Biopharma, Rheos Medicines, Seres Therapeutics, Shire, Synergy Pharmaceuticals, Takeda, Target PharmaSolutions, Theravance Biopharma R&D, TiGenix, Vivelix Pharmaceuticals; honoraria for speaking in CME programs from Takeda, Janssen, Lilly, Gilead, Pfizer, Genetech; research funding from Celgene, Pfizer, Takeda, Theravance Biopharma R&D, Janssen.

CR has received consulting fees from Alimentiv, Inc. (formerly Robarts Clinical Trials, Inc.).

MCD has received consulting fees for Abbvie, Arena, Boehringer Ingelheim, Bristol-Myers Squibb, Celgene, Eli Lilly, Janssen, Pfizer, Prometheus Biosciences, Takeda, Target RW; grant support from Abbvie, Pfizer, Janssen, Co-Founder Trellus Health, Mi-Test Health and Cornerstones Health.

FR is consultant to Agomab, Allergan, AbbVie, Boehringer-Ingelheim, Celgene, Cowen, Genentech, Gilead, Gossamer, Guidepoint, Helmsley, Index Pharma, Jannsen, Koutif, Metacrine, Morphic, Origo, Pfizer, Pliant, Prometheus Biosciences, Receptos, RedX, Roche, Samsung, Takeda, Techlab, Thetis, UCB and receives funding from the Crohn's and Colitis Foundation of America, the Helmsley Charitable Trust, Kenneth Rainin Foundation and the National Institute of Health.

NH has received consulting fees from Abbvie, Celgene, and Lilly USA and has had service contracts with Abbvie and Celgene.

MTA has consulted for Janssen, Prometheus Bioscience, Takeda, Focus Medical Communications, Pfizer, Boehringer Ingelheim Pharmaceuticals, Gilead, Imedex, Cornerstone Health, Inc, Landos Biophama, UCB Biopharma SRL, Eli Lilly, and Cosmo Biopharma. MTA has received grant support from Prometheus Bioscience, Takeda, and Pfizer.

RVB has received Grant/Research support/Speaker fees (all paid to employer for research support): AbbVie, Ferring, Janssen, Shire, Takeda, Emerge Health; shareholder in Biomebank.

GYL has received consulting fees from Alimentiv, Inc. (formerly Robarts Clinical Trials, Inc.).

RK has no conflicts of interest to report.

MAV has received consulting fees from Alimentiv, Inc. (formerly Robarts Clinical Trials, Inc.).

EC has no conflicts of interest to report.

WJS reports research grants from Abbvie, Abivax, Arena Pharmaceuticals, Boehringer Ingelheim, Celgene, Genentech, Gilead Sciences, Glaxo Smith Kline, Janssen, Lilly, Pfizer, Prometheus Biosciences, Seres Therapeutics, Shire, Takeda, Theravance Biopharma; consulting fees from Abbvie, Abivax, Admirx, Alfasigma, Alimentiv (Robarts Clinical Trials, owned by Health Academic Research Trust [HART]), Alivio Therapeutics, Allakos, Amgen, Applied Molecular Transport, Arena Pharmaceuticals, Bausch Health (Salix), Beigene, Bellatrix Pharmaceuticals, Boehringer Ingelheim, Boston Pharmaceuticals, Bristol Meyers Squibb, Celgene, Celltrion, Cellularity, Cosmo Pharmaceuticals, Escalier Biosciences, Equillium, Forbion, Genentech/Roche, Gilead Sciences, Glenmark Pharmaceuticals, Gossamer Bio, Immunic (Vital Therapies), Index Pharmaceuticals, Intact Therapeutics, Janssen, Kyverna Therapeutics, Landos Biopharma, Lilly, Oppilan Pharma, Otsuka, Pandion Therapeutics, Pfizer, Progenity, Prometheus Biosciences, Protagonists Therapeutics, Provention Bio, Reistone Biopharma, Seres Therapeutics, Shanghai Pharma Biotherapeutics, Shire, Shoreline Biosciences, Sublimity Therapeutics, Surrozen, Takeda, Theravance Biopharma, Thetis Pharmaceuticals, Tillotts Pharma, UCB, Vendata Biosciences, Ventyx Biosciences, Vimalan Biosciences, Vivelix Pharmaceuticals, Vivreon Biosciences, Zealand

An international consensus to standardize integration of histopathology in ulcerative colitis clinical trials

Christopher Ma, MD, MPH[1,2], **Rocio Sedano, MD**[3], **Ahmed Almradi, MD**[3], **Niels Vande Casteele, PharmD, PhD**[2,4], **Claire E. Parker, MA, MLIS**[2], **Leonardo Guizzetti, MSc, PhD**[2], **David F. Schaeffer, MD**[5], **Robert H. Riddell, MD, FRCPath, FRCPC**[6], **Reetesh K. Pai, MD**[7], **Robert Battat, MD**[8], **Bruce E. Sands, MD MS**[9], **Christophe Rosty, MD, PhD**[10,11], **Marla C. Dubinsky, MD**[12], **Florian Rieder, MD**[13,14], **Noam Harpaz, MD, PhD**[9,15], **Maria T. Abreu, MD**[16], **Robert V. Bryant, MBBS MScR PhD**[17,18], **Gregory Y. Lauwers, MD, FAGPS**[19], **Richard Kirsch, MBChB, PhD**[6], **Mark A. Valasek, MD, PhD**[20], **Eileen Crowley, MB Bch BAO, MRCPI**[21], **William J. Sandborn, MD**[2,4], **Brian G. Feagan, MD**[2,3,22], **Rish K. Pai, MD, PhD**[23], **Vipul Jairath, MD, PhD**[2,3,22]

[1]Division of Gastroenterology and Hepatology, Departments of Medicine and Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

[2]Alimentiv Inc. (formerly Robarts Clinical Trials, Inc.), London, Ontario, Canada

[3]Department of Medicine, Division of Gastroenterology, Western University, London, Ontario, Canada

[4]Division of Gastroenterology, University of California San Diego, La Jolla, California, United States

[5]Department of Pathology and Laboratory Medicine, The University of British Columbia, Vancouver, British Columbia, Canada

[6]Department of Laboratory Medicine & Pathobiology, Mount Sinai Hospital, Lunenfeld-Tanenbaum Research Institute, University of Toronto, Toronto, Ontario, Canada

[7]Division of Anatomic Pathology, Department of Pathology, University of Pittsburgh, Pittsburgh, Pennsylvania, United States

[8]Jill Roberts Center for IBD, Division of Gastroenterology and Hepatology, Weill Cornell Medicine, New York, New York, United States

[9]Department of Medicine, The Dr. Henry D. Janowitz Division of Gastroenterology, Icahn School of Medicine at Mount Sinai Hospital, New York, New York, United States

[10]Faculty of Medicine, The University of Queensland, Brisbane, Queensland, Australia

[11]Envoi Pathology, Brisbane, Queensland, Australia

[12]Department of Pediatrics, Division of Pediatric Gastroenterology, Icahn School of Medicine at Mount Sinai Hospital, New York, New York, United States

[13]Department of Gastroenterology, Hepatology and Nutrition, Digestive Diseases and Surgery Institute, Cleveland Clinic Foundation, Cleveland, Ohio, United States

[14]Department of Inflammation and Immunity, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, Ohio, United States

[15]Department of Pathology, Molecular and Cell-Based Medicine and Division of Gastroenterology, Icahn School of Medicine at Mount Sinai, New York, New York, United States

[16]Crohn's and Colitis Center, Division of Gastroenterology, Department of Medicine, University of Miami Leonard Miller School of Medicine, Miami, Florida, United States

[17]IBD Service, Department of Gastroenterology, The Queen Elizabeth Hospital, Adelaide, South Australia, Australia

[18]Faculty of Health Sciences, School of Medicine, University of Adelaide, Adelaide, South Australia, Australia

[19]DH. Lee Moffitt Cancer Center & Research Institute, and University of South Florida, Tampa, Florida, United States

[20]Department of Pathology, University of California San Diego, La Jolla, California, United States

[21]Division of Pediatric Gastroenterology, Western University, Children's Hospital, London Health Sciences Centre, London, Ontario, Canada

[22]Department of Epidemiology and Biostatistics, Western University, London, Ontario, Canada

[23]Department of Laboratory Medicine and Pathology, Mayo Clinic, Scottsdale, Arizona, United States

Abstract

Background & Aims: Histopathology is an emerging treatment target in ulcerative colitis (UC) clinical trials. We aim to provide guidance on standardizing biopsy collection protocols, identifying optimal evaluative indices, and defining thresholds for histologic response and remission after treatment.

Methods: An international, interdisciplinary expert panel of 19 gastroenterologists and gastrointestinal pathologists was assembled. A modified RAND/University of California Los Angeles appropriateness methodology was used to address relevant issues. A total of 138 statements were derived from a systematic review of the literature and expert opinion. Each statement was anonymously rated as appropriate, uncertain, or inappropriate using a 9-point scale. Survey results were reviewed and discussed prior to a second round of voting.

Results: Histologic measurements collected using a uniform biopsy strategy are important for assessing disease activity and determining therapeutic efficacy in UC clinical trials. Multiple biopsy strategies were deemed acceptable, including segmental biopsies collected according to the endoscopic appearance. Biopsies should be scored for architectural change, lamina propria chronic inflammation, basal plasmacytosis, lamina propria and epithelial neutrophils, epithelial damage, and erosions/ulcerations. The Geboes Score, Robarts Histopathology Index, and Nancy Index were considered appropriate for assessing histologic activity; use of the modified Riley Score and Harpaz Index were uncertain. Histological activity at baseline should be required for enrollment, recognizing this carries operational implications. Achievement of histologic improvement or remission were considered appropriate and realistic therapeutic targets. Current histological indices require validation for pediatric populations.

Conclusions: These recommendations provide a framework for standardized implementation of histopathology in UC trials. Additional work is required to address operational considerations and areas of uncertainty.

SHORT SUMMARY

We convened an interdisciplinary expert panel to create a standardized framework for collecting biopsies, interpreting histologic disease activity, and defining histologic response and remission for UC clinical trials.

Keywords

Geboes Score; histology; inflammatory bowel disease; Nancy Index; Robarts Histopathology Index

INTRODUCTION

Ulcerative colitis (UC) is a chronic inflammatory bowel disease (IBD) that results from a dysregulated mucosal immune response.[1, 2] Historically, medical therapy has been aimed at improving patient symptoms, notwithstanding that achieving clinical remission alone has proved insufficient to change the natural history of UC. The advent of more effective treatments has facilitated the shifting of therapeutic targets towards normalization of objective measures of disease activity using a "treat-to-target" approach.[3] However, the precise target of medical therapy remains contentious since comparison of relevant

treatment targets has not been assessed in a controlled manner. While achieving endoscopic remission is currently recommended as the goal in clinical practice,[3] histopathologic improvement has emerged as a potential treatment target as it is a more sensitive measure of mucosal inflammation. Furthermore, several observational studies have shown that histological remission is associated with lower rates of disease-related complications including hospitalization, corticosteroid use, and colectomy, compared to either resolution of symptoms or endoscopic improvement alone.[4-7] Finally, measuring histologic activity is also conceptually advantageous because the pathophysiology of inflammatory changes in UC begin at the level of the colonic mucosa.

Despite these potential benefits, substantial barriers exist to the use of histopathology as an outcome measure in clinical trials and practice. First, there is an absence of controlled data on which to base decisions. Second, while multiple indices for measuring histologic activity in UC have been developed, there are no widely accepted endpoint definitions, either data-driven or determined by consensus, for histologic response and remission.[8] Third, the operating properties of the different histology indices require further comparative evaluation.[9] Fourth, there is a lack of standardization with respect to appropriate biopsy collection protocols. Specifically, it is uncertain how many biopsies are required or where biopsies should be taken to obtain optimal results. Fifth, there are minimal data to explore the impact of histological remission on long-term clinical outcomes. Despite these uncertainties, the United States Food and Drug Administration has approved a drug label for mucosal healing for ustekinumab using a combined endpoint of endoscopic and histological healing.[10] Recognizing that histology samples are now routinely collected in UC clinical trials and multiple novel therapeutic agents are in the development pipeline, consensus recommendations that address heterogeneity in biopsy acquisition, measurement tools and item definitions for histologic activity, and thresholds for classifying histologic response and remission are urgently needed. The European Crohn's and Colitis Organization recently provided recommendations for histology evaluation in routine clinical practice, however, detailed guidance for clinical trials outside of using validated instruments was missing.[11] Therefore, we assembled an interdisciplinary panel of experts and conducted a multiple round consensus process using modified RAND/University of California Los Angeles (UCLA) appropriateness methodology with the intent of generating recommendations for standardization of histologic assessment in UC clinical trials.

MATERIALS AND METHODS

Systematic literature review

We previously conducted a Cochrane systematic review evaluating histologic scoring systems in UC and specifically updated the review for this exercise.[9] The updated systematic review searched PubMed (1966), EMBASE (Ovid, 1947), MEDLINE (Ovid, 1946), and the Cochrane Library (CENTRAL) up to March 1, 2020 without language restriction. The primary aim was to identify definitions, instruments, and operating properties of histology indices used for the assessment of patients with UC. All study designs (i.e., randomized controlled trials, cohort studies, and case series) were considered for inclusion. The full search terms are summarized in Supplementary Appendix 1. A recursive search of

bibliographies of relevant review articles was also performed, as was a manual review of abstracts submitted to Digestive Disease Week, United European Gastroenterology Week, and the European Crohn's and Colitis Organization Scientific Meeting from 2014 to 2019. Eligible studies evaluated a histology index for measuring disease activity in patients with UC as confirmed by conventional clinical, endoscopic, and histologic criteria. Citations and abstracts were independently screened, and data were extracted by two reviewers (AA and RS). Disagreements were resolved by consensus or arbitration by a separate reviewer (CM).

Histologic scoring data from eligible studies was collected. The reliability, validity, responsiveness, and feasibility of UC histology indices were evaluated. Reliability was assessed by measurement of intra- and inter-rater reliability, test-retest reliability, or internal consistency. Validity was evaluated for content, criterion, and construct validity. Responsiveness was defined by the ability of the index to measure change after a period in which histologic variation could reasonably be expected (e.g., following a treatment of known efficacy). Feasibility was evaluated based on rater assessment of ease of administration and time needed for scoring. A summary of the key results of the systematic review which informed the RAND/UCLA survey are presented narratively.

Expert consensus process

Expert Recruitment—An interdisciplinary, international panel of experts was selected to participate, including ten gastroenterologists and nine pathologists, specializing in IBD, from four countries including the United States, Canada, the Netherlands, and Australia. Participants were selected based upon publication record and expertise in clinical trial design and drug development, clinical epidemiology, or histopathology in UC. The final selection of panelists was performed by CM, RKP, BGF, and VJ. Colorectal surgeons were not included, given that the focus of this consensus was on determining definitions and histology targets in UC clinical trials of medical therapy.

Modified RAND/UCLA Appropriateness Methodology—A modified RAND/UCLA appropriateness methodology that incorporated a Delphi panel approach with iterative rounds of voting and discussion to combine the best available evidence with expert experience was used.[12] The modified RAND/UCLA approach evaluated face validity (i.e., the extent to which an item addresses the concept it purports to measure) and feasibility of items identified in the systematic review, as well as additional items derived from expert panelist opinion.

Two rounds of voting occurred. In the first introductory panel meeting, items identified by the systematic review were summarized, and panelists were given an opportunity to generate additional relevant statements. The complete list of statements was then circulated via an online survey and all panelists anonymously rated each item by appropriateness on a 9-point scale (1=inappropriate, 9=highly appropriate). Each survey item was classified as *inappropriate*, *uncertain*, or *appropriate* based on the median panel rating and degree of panel disagreement, as defined by the RAND/UCLA manual: inappropriate (median score 1 to 3.5 without disagreement); uncertain (median score 3.5 to 6.5 without disagreement or any median score with disagreement); and appropriate (median score 6.5 to 9 without

disagreement). Disagreement was defined as having at least six or more panelists in each of the lowest (1-3) and highest (6-9) three-point regions.

Results of the first-round survey were then collated, distributed to panelists, and reviewed in a moderated teleconference, with the aim of identifying areas of disagreement on item appropriateness and rationale for responses. The survey was then revised based on the panel meeting to improve clarity prior to re-circulation and second round voting. Statement appropriateness for the second round of voting was scored as described above.

RESULTS

Systematic Review

A total of 7776 records were identified. After removing duplicates, 5205 citations were screened, and 192 records underwent full-text review. A total of 113 reports of 40 studies were included (Supplementary Figure 1). The Geboes Score, Robarts Histopathology Index, and Nancy Index were the most studied evaluative indices. Most indices evaluated similar items, including the presence of acute or chronic inflammatory infiltrates, the presence of neutrophils (especially in the lamina propria or epithelium), and structural changes such as crypt destruction or ulceration. The operating properties of these indices, including the inter- and intra-rater reliability, measures of content, construct, and criterion validity, and responsiveness are summarized Supplementary Tables 1-5.

Item Generation and Survey

Results from the systematic review were used to inform the survey statements. Items were grouped according to the following topics: standardization of biopsy acquisition, histologic items/indices for measurement of disease activity in UC clinical trials, configuration of clinical trial endpoints in UC, and considerations for pediatric and adolescent participants with UC. The first survey consisted of 125 items. After a moderated teleconference to review the results, a final survey consisting of 138 items was circulated. Overall, 100 (72%) items were considered appropriate, 35 (25%) uncertain, and 3 (2%) inappropriate. Key items, including panelist ratings, are summarized in Tables 1 to 3.

Appropriateness of Items

General Considerations and Standardization of Biopsy Acquisition—The panel determined that histologic measurements, collected using a uniform biopsy strategy, are important for assessment of disease activity and determination of therapeutic efficacy in UC clinical trials (Table 1). Although the panel recognized that there may be cases where endoscopic and histologic disease activity is discordant, it was determined that endoscopic appearance of the mucosa should guide the site of biopsy procurement. Biopsies should be taken from the ulcer edge if ulcer(s) are present, or from the most abnormal area in macroscopically inflamed mucosa without ulceration, or randomly from endoscopically normal mucosa. Standard forceps are appropriate for biopsy acquisition.

At screening, segmental biopsies taken from the rectum, sigmoid, descending, transverse, and ascending colon were deemed to be appropriate if a colonoscopy was performed;

whereas three segment biopsies from the descending, sigmoid colon and rectum are appropriate for sigmoidoscopy. Some panelists felt this strategy may not be feasible given operational implications for histopathology reading volume. Alternative strategies of taking biopsies from the worst affected area at either 0 to 25 cm or 15 to 25 cm from the anal verge were also considered appropriate. However, panelists acknowledged that the strategy of using a specific distance is: 1) subject to measurement error due to looping of the endoscope; 2) may exclude evaluation of the rectum; 3) discordant with the endoscopic evaluation that accounts for all of the mucosa; 4) problematic because the healing of UC after medical treatment may be patchy; and 5) limited sampling may be susceptible to confounding by non-UC pathology such as ischemia or segmental colitis associated with diverticular disease.

For induction trials, the panelists deemed biopsy procurement at 8 to 12 weeks after randomization to be appropriate, whereas there was uncertainty between 14 to 20 weeks. In maintenance trials, biopsy procurement at week 52 relative to randomization was deemed to be appropriate. At follow-up, the panelists determined that biopsies should still be taken from the worst affected area in each segment, even if that area was different compared to baseline, although some panelists endorsed the importance of obtaining biopsies from both the worst affected area as well as the same area sampled at baseline.

Biopsies should be fixed in formalin for a minimum of 6 hours and maximum of 36 hours to ensure optimal tissue samples for routine hematoxylin and eosin (H&E) staining and subsequent molecular and immunohistochemistry (IHC) studies. The panelists identified proper orientation of the biopsies in the tissue block after procurement as a necessary step for accurate scoring, such that the long axis of the colonic crypts is visualized in the tissue section. The panel voted that H&E staining is sufficient for measuring disease activity. Although it was recognized that IHC can provide valuable additional information, there was uncertainty as to whether IHC should be routinely required to measure disease activity. Specifically, it was uncertain if myeloperoxidase (MPO) expression by IHC should be assessed in UC clinical trials as it would add time and expense and could potentially overestimate histologic activity as MPO is not expressed by neutrophils only.

Histologic Items and Indices to Measure Disease Activity in UC Clinical Trials

—Histologic items deemed appropriate for disease measurement included: degree of architectural change/distortion, lamina propria chronic inflammation (lymphocytes and plasma cells), basal plasmacytosis, lamina propria neutrophils, epithelial neutrophils, epithelial damage (surface epithelial injury and crypt destruction), and presence of erosions and ulcerations (Table 2). When evaluating chronic inflammatory infiltrate, the panel determined that both basal plasmacytosis and lamina propria cellularity should be assessed. There was uncertainty regarding the appropriateness of quantifying and measuring lamina propria eosinophils. For each item to be measured, the panel determined it was appropriate for the biopsy fragment generating the worst score for that item to be scored, rather than basing the score on the average involvement across all fragments.

Several histology indices were considered. The Geboes Score was considered appropriate for assessing histologic activity in UC and classifying active disease, although the panel

acknowledged that it was not specifically designed for measuring change in histologic activity due to the hierarchical scoring. Similarly, panelists were uncertain if converting the Geboes Score to a continuous 0 to 22 scale was appropriate for this purpose. The panelists were asked to evaluate both the original descriptors of the Geboes Score as well as an alternate definition adapted to use more quantitative cut-offs (Supplementary Table 6). Panelists rated both the original and adapted definitions as appropriate. The Robarts Histopathology Index and the Nancy Index were determined to be appropriate for assessing histologic disease activity in UC, classifying active versus inactive disease, and measuring change in histologic disease activity after treatment. The panelists were uncertain if the modified Riley Score or the Harpaz Index are appropriate instruments to use in UC clinical trials.

Configuration of Clinical Trial Endpoints and Outcome Definitions—There was agreement that a minimum degree of histological activity at baseline as defined by a Geboes Score of 3.1 (minimum of neutrophilic inflammation in the epithelium), a Robarts Histopathology Index score of 4, or Nancy Index score of 2 should be required for enrollment in clinical trials; despite the fact that this is not routinely performed in current trials (Table 3). While this recommendation was deemed appropriate for trials of mild-to-moderate UC (in order to rigorously confirm active inflammation), there was uncertainty if minimum histological activity should also be required in patients with moderate-to-severe disease.

Achievement of histologic remission or histologic improvement, or measurement of change in histology scores were all voted to be realistically achievable and appropriate endpoints for use in clinical trials. There was agreement that histologic remission should be defined by the absence of neutrophilic inflammation in the mucosa, both at the end of induction and maintenance therapy. The panel voted that a 7-point reduction in the Robarts Histopathology Index from baseline or 1-point reduction in the Nancy Index from baseline were the most appropriate definitions for histologic improvement. Histologic improvement as defined by a Geboes Score of 3.1 was also considered appropriate. The definition of mucosal healing based on endoscopic improvement and histologic remission was determined to be appropriate, in accordance with recommendations from regulatory agencies. The use of histology as either a primary or secondary endpoint in UC clinical trials was voted as appropriate, although we acknowledge that the use of histology as a primary endpoint may not be acceptable to regulatory authorities. However, it was uncertain whether direct comparison of the biopsies should be performed when assessing change in disease activity between time points for a given subject if the pathologist is blinded to timepoint and treatment assignment.

Several histologic treatment targets were considered. Histologic remission, defined as the absence of neutrophilic inflammation in the mucosa, with or without normal to mild increase in lamina propria chronic inflammation, was voted to be an appropriate and realistic histologic target in UC. However, there was uncertainty if it was realistic to achieve improvement in lamina propria chronic inflammation at the end of induction therapy. The panel was also uncertain if additionally targeting a normal to mild increase in lamina propria eosinophils for histologic remission is appropriate or realistic. There was uncertainty

if complete histologic normalization is an appropriate therapeutic target as there may be residual chronic architectural distortion, and there was agreement that is not a realistic target based upon the efficacy of current therapies.

Considerations for Pediatric and Adolescent UC Patients—While the panel agreed that histological disease indices developed for adult patients are currently appropriate to use in adolescent and pediatric patients, there was recognition that there are potentially different pathological features, including more patchy histologic involvement, more frequent pancolonic disease extent, and potentially different patterns of healing in pediatric compared to adult patients. The panel identified greater sampling variation as a potential confounder of thresholds for histologic outcomes in children, and a need to validate histologic indices in this population.

DISCUSSION

Histopathology is an increasingly important modality for assessing disease activity in patients with UC. As most drug development programs routinely incorporate biopsy collection at enrollment and endpoint evaluation, the development of a standardized approach to assessing histopathology in clinical trials is needed. This expert panel developed a framework for implementing histopathology in UC trials, including recommendations for standardization of biopsy procurement and handling protocols, histologic items, and indices appropriate for measuring disease activity, and configuration of histopathology endpoints and outcome definitions.

Panelists agreed that a uniform biopsy strategy is needed to optimally measure disease activity. Historically in UC trials, two biopsies collected from the area of worst inflammation ~15 to 25 cm from the anal verge were used to measure histologic endpoints.[13] However, several potential problems with this paradigm were identified: 1) patients with UC may have patchy healing after treatment; 2) it is unclear if distal biopsies adequately reflect proximal inflammation in patients with more extensive disease; 3) this convention can result in biopsy sampling from an area that does not correlate with endoscopic assessment, which is scored based on the most inflamed segment observed[14, 15]; and 4) this strategy ignores inflammation in the rectum, which is typically the last segment in the colon to heal. This discrepancy may be further compounded as some trial programs require full colonoscopy rather than sigmoidoscopy alone. Notably, the panel voted that taking either 3- or 5-segment biopsies (depending on the extent of the procedure) as appropriate. While this strategy would better capture heterogeneity in histologic disease activity, there are practical implications to taking segmental biopsies both at baseline and follow-up, including increased operational burdens for proper biopsy labeling and handling, slide preparation and fixation, central reader pathologist assessment, limitation to procurement of additional biopsies for biomarker analyses and procedural burden to patients.

Panelists agreed that, at follow-up, it is appropriate to target biopsies at the same area as baseline procurement, in addition to sampling the area of worst endoscopic inflammation, accepting that these may be different locations. It should be stressed that such an approach risks bias towards finding no difference in histological inflammation between baseline and

follow-up. Additionally, there are practical limitations to identifying the previous biopsy site, and although generally considered safe, it was agreed that endoscopic tattooing was inappropriate for this purpose.[16] Two to four biopsies per segment were considered appropriate, and this is consistent with a recent study demonstrating that blinded evaluation of two or three biopsies could reliably assess histologic disease activity in a single colonic segment using the Robarts Histopathology Index.[17] When multiple biopsies are taken, the panel determined each item should be scored based on the fragment that exhibits the worst score for each item, rather than averaging scores across fragments. This recommendation was primarily driven by considerations of maximizing responsiveness within a clinical trial to identify potentially efficacious agents. In a situation analogous to central endoscopic assessment based on the worst mucosal appearance, which has resulted in low endoscopic placebo rates,[18] there were concerns that averaging histology scores across fragments may result in smaller treatment effect sizes, especially if biopsies are taken from areas of normal mucosa.

The panel determined that a minimum degree of histologic activity should be required at baseline for trial enrollment (Geboes Score ≥ 3.1, Robarts Histopathology Index ≥ 4, Nancy Index ≥ 2). To date, this practice has not been routinely implemented. However, it was ultimately considered that if histologic response or remission is to be adopted as a primary or secondary trial endpoint, patients should be required to have microscopic disease activity at baseline to meaningfully report and interpret these outcomes. The panel recognized that requiring baseline histology for enrollment prior to randomization would have operational implications as biopsy samples would need to be processed and centrally read within a reasonable turnaround time after the screening colonoscopy. Alternatively, patients found to lack baseline histopathologic inflammation could be excluded post-randomization in a pre-specified modified intent-to-treat approach.

A second concern was that imposing histologic requirements at baseline could further restrict enrollment in UC trials. Given the discordance between endoscopic and histologic findings in patients with UC,[19] it is plausible that some patients who meet all other criteria for enrollment would be excluded solely on the basis of absent histologic inflammation (which may be related to sampling error alone). It was estimated from the personal experience of the panellists that this situation could arise in approximately 10% of cases.[19-21] Hence, there was substantial discussion on the appropriate minimum threshold and the panel was uncertain if baseline histology should be required to enroll patients with otherwise moderate-to-severe disease. Selecting a higher threshold for enrolment may be too restrictive, and therefore, a minimum of having at least mild neutrophilic inflammation was chosen by the panel, recognizing that many patients will have more significant histologic disease. There may also be a subset of patients with very mild histologic activity, who may not have sufficient baseline disease to meet the suggested definitions of histologic response. In these instances, reporting outcomes using continuous measures (e.g. mean change in the Robarts Histopathology Index) should be considered. Nevertheless, qualifying patients using histology as the primary entry criterion may facilitate participation in trials of patients with mild-to-moderate UC, where the most common reason for screening failure is minimal endoscopic inflammation and may additionally provide confidence that symptoms are due to

UC rather than functional overlap. Additional research is required to understand the effect of different baseline cut-offs of histologic activity on trial recruitment and readouts.

We included only one statement in this exercise relating to the definition of 'mucosal healing', which aimed to capture the current US regulatory recommendations that both endoscopic and histologic disease activity should be assessed for labeling purposes. However, the meaning of 'mucosal healing' has evolved and become ambiguous; in clinical practice, it typically refers to endoscopic improvement (Mayo endoscopic subscore [MES]=1) or endoscopic remission (MES=0) alone without consideration of histopathology. Treatment effect sizes have been shown to vary when considering endoscopic vs. histologic outcomes at the same timepoint.[22] Therefore, we argue that 'mucosal healing' is an insufficiently precise term to describe therapeutic response, both endoscopic and histologic outcomes should be reported separately, and a more accurate descriptor such as "combined endoscopic-histologic remission" or "histo-endoscopic remission" should be used when reporting the composite endpoint. Understanding whether achieving endoscopic remission is sufficient or whether targeting histologic remission will result in better short and long-term outcomes is an important research priority, and a large multicenter clinical trial randomizing patients to different therapeutic endpoints is currently recruiting (NCT04259138).

There was agreement that the presence of neutrophilic infiltration is the most important measure of histologic disease activity in UC. Conversely, the importance of measuring eosinophils was less clear. Although some studies have suggested that mucosal eosinophilia is a negative predictor of treatment response, the panel did not believe that there is sufficient evidence to routinely require eosinophil measurement at this time.[23, 24] While a simpler definition of histologic remission based on the absence or presence of neutrophils was considered, the panel discussed that a definition of histologic remission requiring complete absence of neutrophilic inflammation may be too stringent to achieve with current therapies, and that standardizing histologic assessment using formalized scoring tools is an important priority for both clinical trials and in clinical practice. The Geboes Score, Robarts Histopathology Index, and Nancy Index were considered appropriate instruments for assessing disease activity and have been the most thoroughly evaluated instruments in the literature. There are strengths and limitations to each scoring system and direct comparisons of the operating properties between instruments remains a research priority. It was recognized that the Geboes Score was originally designed as a classification scheme and empirically converting it to either a 6-point or 22-point ordinal or continuous measure based on the highest subscore needs validation. Although there was discussion that the wider dynamic range of the Robarts Histopathology Index may provide a theoretical advantage for measuring responsiveness compared to the Nancy Index, a post-hoc analysis of biopsies collected in the TOUCHSTONE ozanimod trial by four blinded pathologists demonstrated that the standardized effect size was similar across the Geboes Score, Robarts Histopathology Index, Nancy Index, and modified Riley Score.[25]

Multiple definitions of histologic improvement and remission were considered by the panel. In the phase III UNIFI program evaluating ustekinumab in patients with moderate-to-severe UC, histologic improvement was defined as the composite of neutrophil infiltration in < 5% of crypts, with no evidence of crypt destruction, erosions, ulcerations, or granulation tissue.

Achieving this outcome was significantly associated with clinical remission and improved disease activity scores at week 44.[22] The corresponding Geboes Score (3.1) was voted as an appropriate definition of histologic improvement by the panel. For histologic remission, a stricter definition of complete elimination of mucosal neutrophils (Geboes < 2B.1 or Nancy Index < 2) was voted as appropriate, which is consistent with several studies illustrating the clinical benefits of achieving this target.[26] However, there was less clarity on whether chronic inflammatory changes in the lamina propria and architectural distortion should be used in defining histologic targets for treatment. Recent studies have reported that patients who achieve complete histologic normalization are at a reduced risk of clinical relapse compared to patients without normalization.[6, 27] Furthermore, Cushing *et al.* observed that both architectural changes and chronic inflammatory infiltrate were predictive of disease relapse within two years.[27] However, the panel was unclear whether this is appropriate or realistic to achieve with currently available therapies, particularly within the timeframe of an 8- to 12-week induction trial.

Our study has several strengths. We included internationally recognized IBD gastroenterologists and pathologists, with extensive experience in trial design and histopathology interpretation. Additionally, we used rigorous methodology to combine the best available evidence from the literature with the clinical expertise of the panel. Finally, we addressed a broad range of issues pertaining to implementing histopathology in clinical trials from biopsy acquisition to endpoint evaluation. However, we acknowledge some important limitations. First, there were several items for which voting was primarily based on expert opinion as there was limited empirical evidence to guide decision making. We have highlighted these items throughout the discussion as areas of research priority. Second, the modified RAND/UCLA method is not designed to force consensus. Therefore, there are some items voted as appropriate which may seem contradictory to or overlapping with other statements. For example, both segmental sampling and sampling from the worst affected area 0 to 25 cm from the anal verge were rated as appropriate strategies. These statements generally reflect topics of discrepancy where additional data are required to determine the *most* appropriate method. Third, there was a high volume of complex statements given the breadth of the scope of this initiative, and this may have potentially contributed to panelist fatigue.

In summary, we have developed a framework for appropriately integrating histopathology assessment in UC clinical trials through an expert panel. Key conclusions include the importance of assessing histopathology using a standardized biopsy acquisition and handling protocol, measuring histologic activity using validated instruments at both enrollment and endpoint adjudication, and incorporating histologic readouts in clinical trial endpoints in accordance with regulatory guidance. These results can be used by sponsors and regulators to help inform biopsy procurement strategies relevant to their drug development program. We also highlight several research priorities. Additional studies are required to: 1) understand the relative performance characteristics of the different histology indices; 2) compare treatment effects and explain the discrepancies between endoscopic and histologic measurements of disease activity; 3) delineate the prognostic implications of varying thresholds for defining histologic response and remission; 4) recognise the effect on trial recruitment and readouts when minimum standards of histologic activity are required

for enrolment; 5) determine whether achieving histological plus endoscopic remission is casually related to superior outcomes compared to endoscopic remission alone; and 6) maximize the efficiency of histopathology evaluation to facilitate the higher burden of biopsy sampling that has been recommended.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ABBREVIATIONS:

IHC	immunohistochemistry
IL	Interleukin
JAK	Janus Kinase
MPO	Myeloperoxidase
TNF	Tumour Necrosis Factor

REFERENCES

1. Kim DH, Cheon JH. Pathogenesis of Inflammatory Bowel Disease and Recent Advances in Biologic Therapies. Immune Netw 2017;17:25–40. [PubMed: 28261018]

2. Feuerstein JD, Moss AC, Farraye FA. Ulcerative Colitis. Mayo Clin Proc 2019;94:1357–1373. [PubMed: 31272578]

3. Peyrin-Biroulet L, Sandborn W, Sands BE, et al. Selecting therapeutic targets in inflammatory bowel disease (STRIDE): determining therapeutic goals for treat-to-target. Am J Gastroenterol 2015;110:1324–1338. [PubMed: 26303131]

4. Pai RK, Jairath V, Vande Casteele N, et al. The emerging role of histologic disease activity assessment in ulcerative colitis. Gastrointest Endosc 2018;88:887–898. [PubMed: 30142351]

5. Bryant RV, Burger DC, Delo J, et al. Beyond endoscopic mucosal healing in UC: histological remission better predicts corticosteroid use and hospitalisation over 6 years of follow-up. Gut 2016;65:408–14. [PubMed: 25986946]

6. Christensen B, Hanauer SB, Erlich J, et al. Histologic Normalization Occurs in Ulcerative Colitis and Is Associated With Improved Clinical Outcomes. Clin Gastroenterol Hepatol 2017;15:1557–1564 e1. [PubMed: 28238954]

7. Pai RK, Hartman DJ, Rivers CR, et al. Complete resolution of mucosal neutrophils associates with improved long-term clinical outcomes of patients with ulcerative colitis. Clin Gastroenterol Hepatol 2020;18:2510–2517. [PubMed: 31843598]

8. Ma C, Panaccione R, Fedorak RN, et al. Heterogeneity in definitions of endpoints for clinical trials of ulcerative colitis: A systematic review for development of a core outcome set. Clin Gastroenterol Hepatol 2018;16:637–647. [PubMed: 28843356]

9. Mosli MH, Parker CE, Nelson SA, et al. Histologic scoring indices for evaluation of disease activity in ulcerative colitis. Cochrane Database Syst Rev 2017;5:CD011256. [PubMed: 28542712]

10. Sands BE, Sandborn WJ, Panaccione R, et al. Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis. N Engl J Med 2019;381:1201–1214. [PubMed: 31553833]

11. Magro F, Doherty G, Peyrin-Biroulet L, et al. ECCO position paper: Harmonisation of the approach to ulcerative colitis histopathology. J Crohns Colitis 2020;2020 Jun 6:1–9. [PubMed: 31605526]

12. Fitch K, Berstein SJ, Aguilar MD, et al. The RAND/UCLA appropriateness method user's manual. Santa Monica, CA: RAND, 2001.

13. Ma C, Hussein IM, Al-Abbar YJ, et al. Heterogeneity in definitions of efficacy and safety endpoints for clinical trials of Crohn's disease: A systematic review. Clin Gastroenterol Hepatol 2018;16:1407–1419.e22. [PubMed: 29596987]

14. Seow CH, Romanova A, Irwin SP, et al. Infliximab induces mucosal healing in a proximal to distal direction in acute ulcerative colitis. Gastroenterology 2011;140:S62–S63.

15. Kim B, Barnett JL, Kleer CG, et al. Endoscopic and histological patchiness in treated ulcerative colitis. Am J Gastroenterol 1999;94:3258–3262. [PubMed: 10566726]

16. Committee AT, Kethu SR, Banerjee S, et al. Endoscopic tattooing. Gastrointest Endosc 2010;72:681–685. [PubMed: 20883844]

17. Battat R, Vande Casteele N, Pai RK, et al. Evaluating the optimum number of biopsies to assess histological inflammation in ulcerative colitis: a retrospective cohort study. Aliment Pharmacol Ther 2020;52:1574–1582. [PubMed: 32981088]

18. Ma C, Guizzetti L, Panaccione R, et al. Systematic review with meta-analysis: endoscopic and histologic placebo rates in induction and maintenance trials of ulcerative colitis. Aliment Pharmacol Ther 2018;47:1578–1596. [PubMed: 29696670]

19. Colombel JF, Keir ME, Scherl A, et al. Discrepancies between patient-reported outcomes, and endoscopic and histological appearance in UC. Gut 2017;66:2063–2068. [PubMed: 27590995]

20. Shah J, Dutta U, Das A, et al. Relationship between Mayo endoscopic score and histological scores in ulcerative colitis: A prospective study. JGH Open 2020;4:382–386. [PubMed: 32514441]

21. Irani NR, Wang LM, Collins GS, et al. Correlation between endoscopic and histological activity in ulcerative colitis using validated indices. J Crohns Colitis 2018;12:1151–1157. [PubMed: 29893824]

22. Li K, Marano C, Zhang H, et al. Relationship between combined histologic and endoscopic endpoints and efficacy of ustekinumab treatment in patients with ulcerative colitis. Gastroenterology 2020;2020 Aug 24:S0016-5085(20)35107-6.

23. Kim EM, Randall C, Betancourt R, et al. Mucosal eosinophilia is an independent predictor of vedolizumab efficacy in inflammatory bowel diseases. Inflamm Bowel Dis 2020;26:1232–1238. [PubMed: 31633167]

24. Zezos P, Patsiaoura K, Nakos A, et al. Severe eosinophilic infiltration in colonic biopsies predicts patients with ulcerative colitis not responding to medical therapy. Colorectal Dis 2014;16:O420–430. [PubMed: 25040651]

25. Jairath V, Peyrin-Biroulet L, Zou G, et al. Responsiveness of histological disease activity indices in ulcerative colitis: a post hoc analysis using data from the TOUCHSTONE randomised controlled trial. Gut 2019;68:1162–1168. [PubMed: 30076171]

26. Park S, Abdi T, Gentry M, et al. Histological Disease Activity as a Predictor of Clinical Relapse Among Patients With Ulcerative Colitis: Systematic Review and Meta-Analysis. Am J Gastroenterol 2016;111:1692–1701. [PubMed: 27725645]

27. Cushing KC, Tan W, Alpers DH, et al. Complete histologic normalisation is associated with reduced risk of relapse among patients with ulcerative colitis in complete endoscopic remission. Aliment Pharmacol Ther 2020;51:347–355. [PubMed: 31696961]

WHAT YOU NEED TO KNOW

Background & Context:

Assessment of histopathology is essential for defining disease activity in UC. With multiple novel therapies in development, guidance is required to standardize the implementation of histopathology measures in UC clinical trials.

New Findings:

An international, interdisciplinary expert panel was convened to outline appropriate biopsy acquisition protocols, histologic activity indices for disease evaluation, and definitions of histologic outcomes for UC clinical trials.

Limitations:

Some recommendations, such as requiring segmental biopsies and assessment of histologic activity at enrollment, will pose operational challenges for trials that require innovative, efficient solutions.

Impact:

Adopting a standardized framework for biopsy collection and consensus definitions of histologic response and remission will improve the interpretation of these endpoints for novel therapies in UC clinical trials.

Table 1.

Standardization of biopsy acquisition in ulcerative colitis clinical trials.

Item	Median Panel Score (IQR)	Rating
Histologic measurements are important to assess disease activity in UC	9 [8, 9]	Appropriate
Histologic measurements are important to determine therapeutic efficacy in clinical trials of UC	8 [8, 9]	Appropriate
Where to take biopsies		
A uniform biopsy strategy is needed to optimally measure disease activity in UC	9 [8, 9]	Appropriate
The endoscopic appearance of the mucosa should dictate from where biopsies are taken	7 [7, 9]	Appropriate
If an ulcer is present, all biopsies should be taken from the edge of the ulcer	8 [6, 9]	Appropriate
If ulcers are not seen, but there are macroscopically abnormal areas, biopsies should be taken from the most abnormal area	8 [8, 9]	Appropriate
If the endoscopic appearance of the mucosa is normal, random biopsies should be taken	8 [8, 9]	Appropriate
Biopsies should be taken from:		
The worst area in five segments (rectum, sigmoid, descending, transverse, and ascending) at all time points if colonoscopy is performed	7 [7, 8]	Appropriate
The worst area in three segments (rectum, sigmoid, and descending) at all time points if sigmoidoscopy is performed	7 [7, 8]	Appropriate
The worst affected area 15-25cm from the anal verge at all time points	7 [5, 7]	Appropriate
The worst affected area 0-25cm from the anal verge at all time points in order to include the rectum	8 [7, 8]	Appropriate
The worst affected area in the rectosigmoid at all time points	7 [6, 8]	Appropriate
If a certain area was biopsied at baseline, effort should be made to biopsy the same area at subsequent time points (even if the mucosa looks improved or normal)	7 [5, 8]	Appropriate
If a certain area was biopsied at baseline, the biopsies at subsequent time points should be taken from the area of worst endoscopic activity (even if this area is in a different location)	8 [7, 8]	Appropriate
An endoscopic tattoo should be performed to ensure the same area is biopsied at baseline and subsequent timepoints	2 [1, 4]	Inappropriate
The minimum number of biopsies necessary to measure disease activity in UC is:		
1 biopsy per segment/area biopsied	3 [1, 5]	Inappropriate
2 biopsies per segment/area biopsied	7 [6, 8]	Appropriate
3 biopsies per segment/area biopsied	7 [5, 8]	Appropriate
4 biopsies per segment/area biopsied	7 [3, 7]	Appropriate
5 biopsies per segment/area biopsied	3 [2, 5]	Inappropriate
When to take biopsies		
For induction trials, biopsy procurement for histologic assessment of UC should ideally take place in relation to randomization at:		
8 weeks	7 [5, 8]	Appropriate
10 weeks	7 [5, 8]	Appropriate
12 weeks	7 [5, 8]	Appropriate
14 weeks	5 [5, 7]	Uncertain
16 weeks	5 [4, 7]	Uncertain
18 weeks	4 [3, 5]	Uncertain
20 weeks	4 [2, 5]	Uncertain

Item	Median Panel Score (IQR)	Rating
For maintenance trials, biopsy procurement for histologic assessment of UC should take place in relation to randomization at approximately 52 weeks	8 [6, 9]	Appropriate
How to take and prepare biopsies		
Standard biopsy forceps should be used to obtain biopsies	7 [7, 8]	Appropriate
Jumbo biopsy forceps should be used to obtain biopsies	6 [5, 8]	Uncertain
It is acceptable to take biopsies using one bite of the mucosa with one pass of a biopsy forceps	7 [5, 8]	Appropriate
It is acceptable to take biopsies using two bites of the mucosa with one pass of a biopsy forceps	7 [5, 8]	Appropriate
It is acceptable to take biopsies using three bites of the mucosa with one pass of a biopsy forceps	5 [4, 7]	Uncertain
Biopsies should be placed directly in 10% formalin with minimal tissue handling	8 [7, 9]	Appropriate
A minimum and maximum fixation time in formalin should be specified	7 [7, 9]	Appropriate
Biopsies should be fixed for a minimum of 6 hours and maximum of 36 hours to ensure maximum use of the tissue (hematoxylin and eosin, immunohistochemistry, RNA, etc.)	7 [5, 8]	Appropriate
Proper orientation of the biopsies in the tissue block is necessary for accurate scoring	8 [7, 9]	Appropriate
Biopsies should be oriented such that the long axis of the colonic crypts is visualized in the tissue section	8 [7, 9]	Appropriate
Hematoxylin and eosin stained sections are sufficient for histologic evaluation of disease activity	8 [7, 9]	Appropriate
Immunohistochemistry to quantify various cell types should be performed to measure disease activity	4 [3, 5]	Uncertain
Immunohistochemistry can provide valuable additional information regarding disease activity but is not required	7 [4, 8]	Appropriate
Myeloperoxidase expression by immunohistochemistry should be assessed in clinical trials of UC	5 [2, 5]	Uncertain

IQR interquartile range

Table 2.

Histologic items and indices to measure disease activity in ulcerative colitis clinical trials.

Item	Median Panel Score (IQR)	Rating
For each item measured, the score should be based on the biopsy fragment that generates the worst score for that item	9 [7, 9]	Appropriate
For each item measured, the score should be based on the average involvement across all fragments	5 [4, 5]	Uncertain
Histologic items for measuring disease activity		
It is important to measure degree of architectural change/distortion	7 [7, 8]	Appropriate
It is important to measure degree of lamina propria chronic inflammation (lymphocytes and plasma cells)	8 [7, 9]	Appropriate
It is important to measure degree of basal plasmacytosis	8 [6, 9]	Appropriate
It is important to measure degree of lamina propria eosinophils	5 [5, 7]	Uncertain
It is important to measure degree of lamina propria neutrophils	9 [8, 9]	Appropriate
It is important to measure degree of epithelial neutrophils	9 [8, 9]	Appropriate
It is important to measure degree of epithelial damage (including surface epithelial injury and crypt destruction)	9 [8, 9]	Appropriate
It is important to assess for the presence or absence of erosions	9 [8, 9]	Appropriate
It is important to assess for the presence or absence of ulcers	9 [8, 9]	Appropriate
Ulcers and erosions should be distinguished from one another	7 [5, 8]	Appropriate
Histologic indices for measuring disease activity		
The Geboes Score is an appropriate instrument for assessing histologic disease activity in UC	7 [7, 9]	Appropriate
The Geboes Score is an appropriate instrument for classifying histologic disease activity in UC (i.e., distinguishing patients with active vs. inactive disease)	8 [7, 9]	Appropriate
The Geboes Score is an appropriate instrument for measuring change in histologic disease activity in UC	7 [6, 8]	Appropriate
Converting the Geboes Score to a 0-6 scale (taking the highest Geboes subscore and assigning number 0-6) is an appropriate instrument for assessing histologic disease activity in UC	7 [5, 7]	Appropriate
The continuous Geboes Score (where each subscore is given a number from 0-22 and the highest subscore is used to assign the score) is an appropriate instrument for assessing histologic disease activity in UC	6 [5, 7]	Uncertain
The Robarts Histopathologic Index is an appropriate instrument for assessing histologic disease activity in UC	8 [7, 9]	Appropriate
The Robarts Histopathologic Index is an appropriate instrument for classifying histologic disease activity in UC (i.e., distinguishing patients with active vs. inactive disease)	8 [7, 9]	Appropriate
The Robarts Histopathologic Index is an appropriate instrument for measuring change in histologic disease activity in UC	8 [7, 9]	Appropriate
The Nancy Index is an appropriate instrument for assessing histologic disease activity in UC	7 [7, 8]	Appropriate
The Nancy Index is an appropriate instrument for classifying histologic disease activity in UC (i.e., distinguishing patients with active vs. inactive disease)	8 [8, 8]	Appropriate
The Nancy Index is an appropriate instrument for measuring change in histologic disease activity in UC	7 [5, 7]	Appropriate
The modified Riley score is an appropriate instrument for assessing histologic disease activity in UC	6 [5, 7]	Uncertain
The modified Riley score is an appropriate instrument for classifying histologic disease activity in UC (i.e., distinguishing patients with active vs. inactive disease)	6 [5, 8]	Uncertain
The modified Riley score is an appropriate instrument for measuring change in histologic disease activity in UC	6 [5, 7]	Uncertain
The Harpaz Index is an appropriate instrument for assessing histologic disease activity in UC	5 [5, 7]	Uncertain
The Harpaz Index is an appropriate instrument for measuring change in histologic disease activity in UC	5 [5, 7]	Uncertain

Table 3.

Configuration of clinical trial histopathology endpoints and outcome definitions in ulcerative colitis clinical trials

Item	Median Panel Score (IQR)	Rating
Histology outcome configurations and definitions		
Histologic activity in UC is defined by neutrophilic inflammation of the mucosa	8 [7, 9]	Appropriate
Histologic remission in UC is defined by absence of neutrophilic inflammation of the mucosa	8 [7, 9]	Appropriate
Histologic remission is an appropriate histologic endpoint	8 [7, 9]	Appropriate
Histologic remission is a realistic histologic endpoint	7 [7, 8]	Appropriate
The criteria for histologic remission should be the same in trials of mild-to-moderate UC compared to trials of moderate-to-severe UC	7 [7, 9]	Appropriate
Histologic improvement is an appropriate histologic endpoint	8 [7, 9]	Appropriate
Histologic improvement is a realistic histologic endpoint	8 [8, 9]	Appropriate
Histologic improvement is best defined as a Geboes Score 3.1 (< 5% neutrophils in epithelium)	7 [5, 7]	Appropriate
Histologic improvement is best defined as a Robarts Histopathologic Index score 6	7 [5, 7]	Appropriate
Histologic improvement is best defined as a Robarts Histopathologic Index score 9	5 [5, 7]	Uncertain
Histologic improvement is best defined as a 7-point reduction in the Robarts Histopathologic Index score from baseline	6 [5, 7]	Uncertain
Histologic improvement is best defined as a 1-point reduction in the Nancy Score from baseline	7 [6, 7]	Appropriate
Histologic improvement is best defined as a 50% reduction from baseline in the Robarts Histopathologic Index	7 [5, 8]	Appropriate
Histologic improvement is best defined as a decrease (any amount) in the histologic score from baseline	5 [4, 6]	Uncertain
Change in histology score is an appropriate endpoint	8 [7, 9]	Appropriate
Change in histology score is a realistic endpoint	8 [7, 9]	Appropriate
Histology should be part of the primary endpoint in clinical trials of UC	7 [5, 9]	Appropriate
Histology should be a secondary endpoint in clinical trials of UC	8 [7, 9]	Appropriate
Mucosal healing should be defined as endoscopic improvement and histologic remission	8 [7, 9]	Appropriate
When assessing change in disease activity between time points in a given subject, a direct comparison of the biopsies should be performed as long as the pathologist/reader is blinded to timepoint and treatment arm	6 [4, 7]	Uncertain
Baseline histologic disease activity requirements		
A minimum histological disease activity score at baseline should be required for enrollment into all UC clinical trials regardless of disease severity	8 [7, 9]	Appropriate
A minimum histological disease activity score at baseline should be required for enrollment into all UC clinical trials for patients with mild-moderate disease only	7 [5, 8]	Appropriate
A minimum histological disease activity score at baseline should be required for enrollment into all UC clinical trials for patients with moderate-severe disease only	6 [4, 8]	Uncertain
At baseline, there should be at least neutrophilic inflammation of the epithelium (Geboes 3.1)	8 [7, 9]	Appropriate
At baseline, the minimum Robarts Histopathologic Index score should be 4	7 [5, 8]	Appropriate
At baseline, the minimum Nancy Index should be 2	8 [5, 8]	Appropriate
Appropriate and realistic histologic endpoints		
Absence of neutrophilic inflammation in the mucosa (Geboes <2B.1, Nancy <2) is an appropriate histologic endpoint in UC:		
At the end of the induction therapy period	7 [6, 7]	Appropriate

Item	Median Panel Score (IQR)	Rating
At the end of the maintenance therapy period	7 [7, 8]	Appropriate
Absence of neutrophilic inflammation in the mucosa (Geboes <2B.1, Nancy <2) is a realistic histologic endpoint in UC:		
At the end of the induction therapy period	7 [5, 7]	Appropriate
At the end of the maintenance therapy period	7 [7, 8]	Appropriate
Absence of neutrophilic inflammation in the mucosa and normal to only mild increase in lamina propria chronic inflammation (Geboes: 0.0-0.3, 1.0-1.1, 2A.0-2A.3, 2B.0, 3.0, 4.0, 5.0) is an appropriate histologic target:		
At the end of the induction therapy period	6 [5, 7]	Uncertain
At the end of the maintenance therapy period	6 [5, 7]	Uncertain
Absence of neutrophilic inflammation in the mucosa and normal to only mild increase in lamina propria chronic inflammation (Geboes: 0.0-0.3, 1.0-1.1, 2A.0-2A.3, 2B.0, 3.0, 4.0, 5.0) is a realistic histologic target:		
At the end of the induction therapy period	6 [5, 7]	Uncertain
At the end of the maintenance therapy period	6 [6, 7]	Uncertain
Absence of neutrophilic inflammation in the mucosa, normal to only a mild increase in lamina propria chronic inflammation, and normal to only a mild increase in lamina propria eosinophils (Geboes: 0.0-0.3, 1.0-1.1, 2A.0-2A.1, 2B.0, 3.0, 4.0, 5.0) is an appropriate histologic target:		
At the end of the induction therapy period	5 [5, 7]	Uncertain
At the end of the maintenance therapy period	6 [5, 7]	Uncertain
Absence of neutrophilic inflammation in the mucosa, normal to only a mild increase in lamina propria chronic inflammation, and normal to only a mild increase in lamina propria eosinophils (Geboes: 0.0-0.3, 1.0-1.1, 2A.0-2A.1, 2B.0, 3.0, 4.0, 5.0) is a realistic histologic target:		
At the end of the induction therapy period	5 [5, 7]	Uncertain
At the end of the maintenance therapy period	6 [5, 7]	Uncertain
Complete histologic normalization (Geboes: 0.0, 1.0, 2A.0, 2B.0, 3.0, 4.0, 5.0) is an appropriate histologic target		
At the end of the induction therapy period	5 [4, 6]	Uncertain
At the end of the maintenance therapy period	6 [5, 7]	Uncertain
Complete histologic normalization (Geboes: 0.0, 1.0, 2A.0, 2B.0, 3.0, 4.0, 5.0) is a realistic histologic target		
At the end of the induction therapy period.	5 [3, 5]	Uncertain
At the end of the maintenance therapy period.	5 [4, 6]	Uncertain