

A Systematic Review of the Measurement of Endoscopic Healing in Ulcerative Colitis Clinical Trials: Recommendations and Implications for Future Research

Mark A. Samaan, MBBS,^{*,†} Mahmoud H. Mosli, MD,^{*,‡,§} William J. Sandborn, MD,^{*,||}
 Brian G. Feagan, MD,^{*,¶,**} Geert R. D'Haens, MD, PhD,^{*,†} Elena Dubcenco, MD,^{*}
 Kenneth A. Baker, PhD,^{*} and Barrett G. Levesque, MD^{*,||}

Background: Assessment of endoscopic disease activity, as measured by various endoscopic evaluative instruments, is an essential part of quantifying disease activity in clinical trials in patients with ulcerative colitis (UC). Evaluative instruments have specific definitions and operating properties that influence the interpretation of clinical trial results. Our objective was to systematically review all endoscopic evaluative instruments that measure endoscopic disease activity in UC and to describe their definitions and operating characteristics (reliability, responsiveness, and predictive validity).

Methods: We performed a systematic review of evaluative instruments assessing endoscopic disease activity in UC. MEDLINE (Ovid), EMBASE (Ovid), PubMed, the Cochrane Library (CENTRAL), and Digestive Disease Week abstracts of clinical trials were searched from inception to January 2013.

Results: In total, 5885 studies were identified and screened for inclusion criteria. Four hundred twenty-two studies involving 31 evaluative instruments were identified. Two types of indices were found, numerical scoring systems and stepwise grading scales.

Conclusions: Both the endoscopic evaluative instrument selected and the definition chosen for mucosal healing affect the validity of assessing endoscopic disease activity during a clinical trial for UC. Currently, the sigmoidoscopic component of the Mayo Score and the ulcerative colitis endoscopic index of severity show the most promise as reliable evaluative instruments of endoscopic disease activity. However, further validation is required.

(*Inflamm Bowel Dis* 2014;20:1465–1471)

Key Words: ulcerative colitis, endoscopy, indices, clinical trial, disease activity

Ulcerative colitis (UC) is a chronic condition characterized by inflammation of the colonic mucosa.¹ New therapies are able to induce and maintain mucosal healing,^{2,3} which may change the natural history of UC⁴ and reduce the rate of colectomy.⁵ The definition of endoscopic disease activity can have a substantial impact on the operating characteristics of a mucosal healing endpoint in clinical trials, correlation with clinical remission, and prediction of long-term outcomes. This systematic review aims

to identify all the available evaluative instruments used to assess endoscopic disease activity and mucosal healing in UC. We review the development and validation of key evaluative instruments, allowing recommendations to be made regarding optimal assessments of endoscopic activity and mucosal healing for clinical trials. Future research priorities are also identified.

METHODS

Search Strategy

MEDLINE, EMBASE, PubMed, the Cochrane Library (CENTRAL), and Digestive Diseases Week abstracts of clinical trials were electronically searched from their inception to January 16, 2013, for endoscopic evaluative instruments used for the evaluation of UC. A summary of the specific search strategy used is detailed below and a comprehensive description is included in Data, Supplemental Digital Content 1, <http://links.lww.com/IBD/A466>.

Each database was searched for (“ulcerative colitis” OR “inflammatory bowel disease”) AND (“endoscopy” OR “colonoscopy” OR “sigmoidoscopy” OR “proctosigmoidoscopy”) AND (“index” OR “indice” OR “scale” OR “score” OR “grade” OR “Baron” OR “Rachmilewitz” OR “Mayo” OR “Matts” OR “UCEIS” OR “Truelove” OR “Dick” OR “Marks” OR “Feagan” OR “Powell” OR “Lemann” OR “Sutherland”).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.ibdjournal.org).

Received for publication January 31, 2014; Accepted March 13, 2014.

From the ^{*}Robarts Clinical Trials, Inc, London, ON, Canada; [†]Department of Gastroenterology, Academic Medical Center, Amsterdam, the Netherlands; [‡]Department of Medicine, Division of Gastroenterology, King Abdulaziz University, Jeddah, Saudi Arabia; [§]Department of Epidemiology, McGill University, Montreal, QC, Canada; ^{||}Division of Gastroenterology, University of California, San Diego, La Jolla, California; and Departments of [¶]Medicine, and ^{**}Epidemiology and Biostatistics, Western University, London, ON, Canada.

Author disclosures are available in the Acknowledgments.

Reprints: Barrett G. Levesque, MD, Division of Gastroenterology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0956 (e-mail: bglevesque@ucsd.edu).

Copyright © 2014 Crohn's & Colitis Foundation of America, Inc.

DOI 10.1097/MIB.0000000000000046

Published online 14 May 2014.

Study Selection

Two reviewers (M.H.M. and M.A.S.) independently screened citations and abstracts before retrieving full-text publications of all potentially eligible articles. No language restrictions were applied and publications were translated into English where necessary. Study eligibility was then assessed and in cases of disagreement consensus was reached.

RESULTS

Our literature search retrieved a total of 5885 citations. After excluding duplicates (2917), a total of 2968 publications were screened and after removal of 76 animal studies, 2892 remained. After applying eligibility criteria and including an additional 21 articles, a total of 437 articles were analyzed. This comprised 422 studies involving a total of 31 endoscopic scoring systems, as well as 15 related reviews (Fig. 1). The most commonly used and the most recently developed scoring systems are both described in Table 1. The remaining 29 scoring systems identified in our review are described in Tables, Supplemental Digital Contents 2 and 3, <http://links.lww.com/IBD/A467> and <http://links.lww.com/IBD/A468>.

The following summary highlights the historical development, definitions, and operative characteristics of key endoscopic evaluative instruments used in clinical trials for UC.

FIRST ENDOSCOPIC SCORES: BARGEN SCORING AND THE TRUELOVE AND WITTS INDEX

Direct examination of colonic mucosa in assessing patients with colitis was originally described by Bargen in 1937.⁶ His groundbreaking observations were made using a rigid proctoscope in conjunction with a magnifying attachment. This allowed him, for the first time, to describe the mucosal changes seen “in the living patient, from the inception of the disease to its well-advanced state.” Even as part of this fledgling work on mucosal assessment, Bargen clearly recognized the need for stratification of the changes he observed. He described a first stage of disease with “numerous small hemorrhages scattered about the diffusely inflamed mucosa.” A second stage followed with the addition of edematous mucosa, which was described as being “so easily traumatized” by the ridged examining instrument, reflecting friability. Thereafter followed a third and fourth stage as the condition became increasingly fulminant.

In 1955, Truelove and Witts⁷ used the first evaluative instrument for UC endoscopic activity in a clinical trial. In addition to measuring clinical variables to assess disease activity, Truelove and Witts performed serial rigid sigmoidoscopic assessments using a 3-point scale. However, this evaluative instrument did not define the endoscopic descriptors rather it classified patients into 1 of 3 groups: normal or near normal (score of 1), improved (2), no change or worse (3). This rudimentary design allowed for significant interobserver variability.

FIRST VALIDATED ENDOSCOPIC INDEX SCORING SYSTEM FOR UC: THE BARON INDEX

In 1964 Baron et al⁸ evaluated the interobserver variability in describing changes seen in rectosigmoid mucosa. Their study involved 3 observers independently assessing the mucosa of 60 patients with UC. Mucosal examination was performed with a rigid sigmoidoscope and disease activity was rated using a 4-point scale (0–3), largely based on the degree of mucosal friability. This was determined by assessing the degree of mucosal bleeding on brushing the mucosa with a cotton wool pledget. They concluded that continuous variables lead to greater degrees

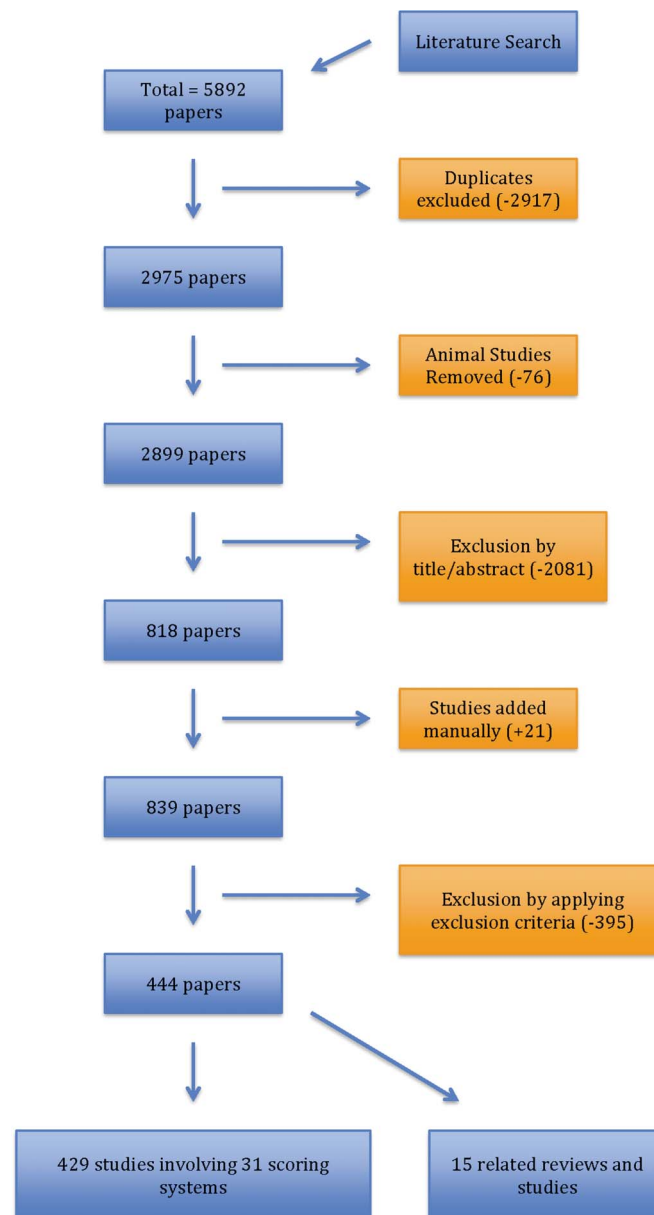


FIGURE 1. Algorithm showing the search strategy used to identify relevant publications based on PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidance.

TABLE 1. Most Commonly Used and Most Recently Developed Scores

Index	Setting	Description of Scale	Extent of Use	Level of Validation
Mayo Score (1987)	Prospective study	Stepwise 4-grade scale (0) Normal or inactive disease (1) Mild disease: erythema, decreased vascular pattern, mild friability (2) Moderate disease: marked erythema, absent vascular pattern, friability, erosions (3) Severe disease: spontaneous bleeding, ulceration	Multiple clinical studies and RCT	Not validated
UCEIS (2012)	To be used in prospective studies. Previously validated in retrospective studies	Numerical grading system generating a total score (3–11) ^a composed of 3 variable subscores a) Vascular pattern (1) Normal (2) Patchy obliteration (3) Obliterated b) Bleeding (1) None (2) Mucosal (3) Luminal—mild (4) Luminal—moderate or severe c) Erosions and ulcers (1) None (2) Erosions (3) Superficial ulcer (4) Deep ulcer	Recently developed and validated— trials using it as an end-point yet to be completed	Partially validated

^aThe grading system has been recently changed from 3–11 to 0–8. The variables of the score remained unchanged. RCT, randomized controlled trial.

of interobserver discrepancy than discontinuous variables that are “capable of close definition.” For example, the agreement for color and granularity were 33% and 40%, respectively. The highest level of agreement was reached for the variable of mucosal “friability,” where it was found that agreement between the observers was approximately 90%. Unfortunately, as this work predates the advent of current methods of statistical analysis, reliability, kappa, and intraclass correlation coefficients were not reported.⁹ The original Baron score itself has never been formally validated to determine its responsiveness and predictive validity.

SUTHERLAND MUCOSAL APPEARANCE ASSESSMENT

As part of a placebo-controlled trial investigating mesal-amine enemas, Sutherland and Martin¹⁰ described a 4-point scale based on serial sigmoidoscopic assessment. The grade (0–3) increased with the degree of mucosal friability. A composite score, the Sutherland Index, was then devised incorporating this scale with clinical variables (stool frequency, rectal bleeding,

mucosal appearance, and physician’s rating of disease activity). This index is also known as the Disease Activity Index and the UC Disease Activity Index. Although the index has been demonstrated to correlate closely with patient-defined remission,¹¹ neither the endoscopic nor the composite score has been validated.

ENDOSCOPIC COMPONENT OF THE MAYO SCORE

Although the studies described above included both clinical and endoscopic assessments of disease activity, the first widely used instrument to incorporate both of these domains into a composite score was the Mayo Score. This is also known as the Mayo Clinic Score (or Index, MCS or MCI), which was described by Schroeder et al in 1987.¹² This instrument takes into account 4 variables: stool frequency, rectal bleeding, a physician’s global assessment, and assessment of changes seen in the rectosigmoid mucosa using a flexible endoscope (Table 1 and Fig. 2). Individual items are rated 0 to 3, giving the composite score a maximum of 12. As with the endoscopic component of the Sutherland index, the Mayo Score is

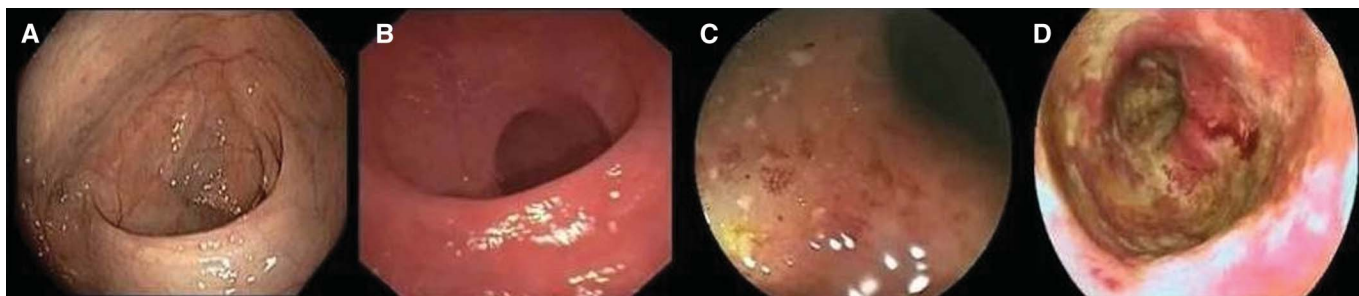


FIGURE 2. Endoscopic images demonstrating increasing grades of inflammation according to the Mayo Score (grade 0–3): (A) 0; (B) 1; (C) 2; (D) 3.

partially based on mucosal friability. It has since been demonstrated that friability is best assessed by the incidental contact of a flexible sigmoidoscope with mucosa rather than the use of a closed biopsy forceps.¹³

The authors suggested definitions for both complete and partial response. However, these definitions relied largely on patient-reported parameters and a physician's global assessment, as the index was originally described as a composite instrument with both endoscopic and clinical components.

Owing to their similar design, some trials have interchanged the endoscopic and clinical components of the Sutherland and Mayo Indices. Studies have successfully demonstrated that improvement in the Mayo Score is a clinically relevant endpoint that correlates with improvement in quality of life measures.¹⁴ In addition, the Active Ulcerative Colitis Trials (ACT-1 and ACT-2) of infliximab showed that mucosal healing, defined as a subscore of 0 or 1 for endoscopy in the Mayo Clinic Score at week 8, was associated with a significantly lower rate of colectomy after 54 weeks ($P = 0.0004$).¹⁵ However, until recent studies no formal validation of the reliability or responsiveness had been carried out. In a placebo-controlled study designed to assess change in disease activity with mesalamine (Asacol; Procter & Gamble, Cincinnati, OH) treatment, a group of experienced central readers scored recorded endoscopies from patients with mild-to-moderate UC before and after treatment. There was excellent intraobserver and interobserver reliability (intraclass correlation coefficient and 95% confidence interval [95% CI]: 0.89 [0.85–0.92] and 0.79 [0.72–0.95], respectively). Initial data from this work suggest that the endoscopic subscore of the MCS is also responsive to change with a treatment of known efficacy. There was a significant difference in the magnitude of change in this endoscopic index between placebo and mesalamine (Asacol; Procter & Gamble) at both 6 and 10 weeks (the median change in endoscopic subscore grades were 0.29 and 0.52, respectively, 2-sample t test, $P = 0.017$ and <0.001).¹⁶ Despite this, many questions remain unanswered: where exactly should the disease severity be scored? What is the minimal insertion length? How should focal healing be handled?

UC ENDOSCOPIC INDEX OF SEVERITY

In 2012, Travis et al¹⁷ started with the parameters included in the Baron score and by using a combination of regression

techniques and central reading of recorded endoscopy they developed the UC Endoscopic Index of Severity (UCEIS). As part of this reevaluation they found that friability and mucosal hemorrhage had similar reliability (weighted inter-investigator kappa values of 0.40 and 0.37, respectively). Although the 2 could not be differentiated on statistical grounds, the latter was a better compliment to the remaining components of the instrument. It has less overlap with the discriminative features of the other parameters and therefore allowed for a more comprehensive description of the observed endoscopic disease activity. The weighted interinvestigator kappa for erosion/ulceration and vascular pattern were both 0.42. Using regression modeling, a 3-component index with a total score of 3 to 11 was created based on vascular pattern (scored 1–3), bleeding (1–4), and erosion/ulceration (1–4) (Fig. 3). It was demonstrated to have high interobserver reproducibility and that 90% (pR^2 90%) of the variance when assessing overall endoscopic severity could be captured using this method. The index was also shown to be a good predictor of overall severity when compared with mean overall severity assessments (pR^2 0.78), judged using a visual analog scales.

Further to this, the process of index validation commenced. In a 2013 study to build on their initial work, Travis et al¹⁸ demonstrated that the UCEIS accounted for a mean of 88% of the variability in overall endoscopic severity. Furthermore, satisfactory intraobserver and interobserver reliability was observed. Intrainvestigator agreement ranged from moderate to very good for the descriptors individually (reliability ratios ranging from 0.47 for bleeding to 0.87 for vascular pattern). Good intrainvestigator determination of the overall UCEIS score (weighted kappa 0.72 [95% CI, 0.61–0.82]) was also demonstrated. Moderate interinvestigator agreement was seen for each descriptor as well as the score as a whole (weighted kappa 0.50 [95% CI, 0.49–0.52]).¹⁹ The UCEIS descriptors were also shown to substantially correlate with a global rating of endoscopic severity (median Pearson correlation coefficient between UCEIS and visual analog scale, 0.93). In addition, Feagan et al²⁰ showed excellent intrareader and interreader agreement for the UCEIS: intraclass correlation coefficients = 0.89 (95% CI, 0.85–0.93) and 0.83 (95% CI, 0.77–0.88), respectively, in a study of 7 experienced central readers. It should be noted that as part of the validation process, the authors decided to rework the scores assigned to individual parameters so that 0,

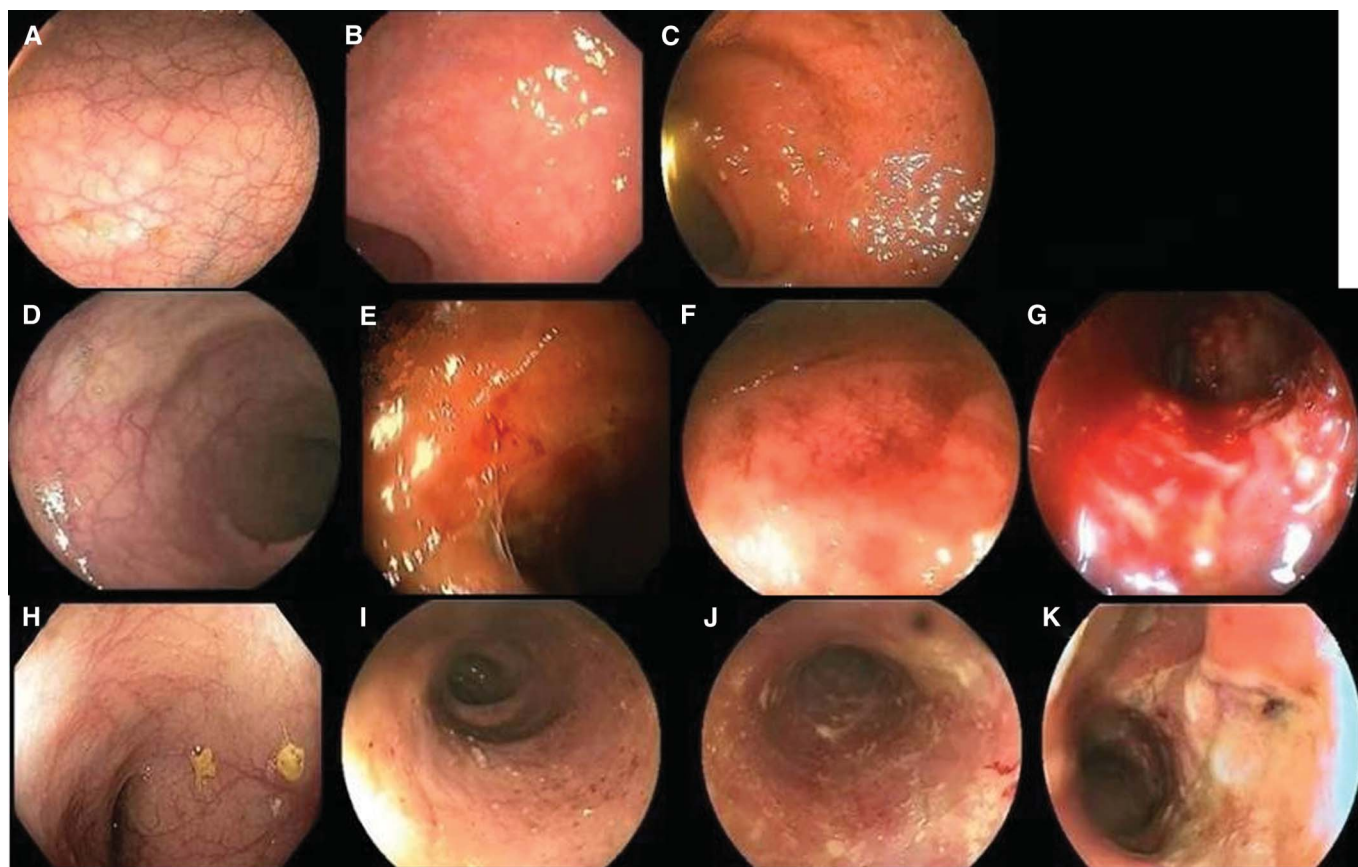


FIGURE 3. Endoscopic images demonstrating the UCEIS scoring system for: vascular pattern (1–3): (A) 1, (B) 2, and (C) 3; bleeding (1–4): (D) 1, (E) 2, (F) 3, and (G) 4; erosions and ulcerations (1–4): (H) 1, (I) 2, (J) 3, and (K) 4.

rather than 1, denotes normality. This gives the index a range of 0 to 8 instead of 3 to 11 but the descriptors of each parameter remain unchanged (Table 1).

UC COLONOSCOPIC INDEX OF SEVERITY

Whether the severity of mucosal inflammation in UC can be characterized by sigmoidoscopic examination is a matter of active debate. Limited data have suggested that the healing in UC can occur in a patchy manner.^{21,22} Samuel et al^{23,24} recently developed and partially validated a colonoscopic scoring system, which grades mucosal changes throughout the entire colon. The UC Colonoscopic Index of Severity (UCCIS) comprises 5 components: vascular pattern, granularity, friability, ulceration, and global severity of damage. The reliability of these variables was evaluated using a library of 50 UC colonoscopy videos, which were examined by 8 experienced central readers, who scored each segment of the colon. Of the parameters investigated, all showed good to excellent interobserver agreement except for friability. To further validate the score, the authors also demonstrated a moderate correlation between the UCCIS and laboratory markers of disease (C-reactive protein [$P < 0.001$], albumin [$P < 0.001$], and hemoglobin [$P < 0.01$]) and a good correlation with patient-defined remission ($P < 0.01$). Despite these validation steps,

significant questions regarding the feasibility of the UCCIS and the segmental responsiveness remain unanswered. Feasibility issues center on the fact that examining the entire colon, rather than just the distal portion, is less tolerable (requiring oral bowel preparation) and more costly.

CENTRAL READING OF ENDOSCOPIC SCORING

Central reading of endoscopic inflammation in UC trials is a relatively novel strategy, first pioneered in a 2009 study of delayed-release mesalamine in moderately active UC conducted by Sandborn et al.¹³ Owing to the experience of the central reader and the fact they remain blinded of the treatment assignments, this method has been shown to reduce inclusion bias and placebo rates in clinical trials. In a recent study, Feagan et al²⁰ demonstrate the advantages of a central reading system by studying the induction of remission using mesalamine in symptomatic UC patients. Inclusion to the study required a Mayo endoscopic subscore of 2 or greater. Two hundred eighty-one patients were first assessed by site investigators and then centrally by a single expert reader. Through this process, 31% of patients said to meet the inclusion criteria by site investigators were subsequently deemed ineligible by the central reader. In a post hoc analysis, these “ineligible”

patients were excluded. By comparing intention-to-treat results, which included all randomized patients, and the post hoc analyses in which the ineligible patients were excluded, the authors were able to show a reduction in the placebo response rates and a subsequent increase in the estimate of treatment effect. It is widely recognized that high placebo response rates are an important factor in the negative outcome of some clinical trials.²⁵ In light of these findings, there is optimism that central reading could play an important role in evaluating the efficacy of new therapies by objectively providing reliable and valid evaluative endpoints with clear operating characteristics. Many clinical trials, which use central reading, are now underway.

FUTURE VALIDATION NEEDS, IDEAL EVALUATIVE INSTRUMENTS, PROOF-OF-CONCEPT TRIALS, CONTROVERSY, AND REGULATORY IMPLICATIONS

Therapy for UC targets mucosal inflammation defined by evaluative instruments. Reliable, valid, responsive, predictive and feasible evaluative instruments are needed for clinical trials. Clear and definite progress has been made from the advent of direct qualitative visualization of rectal mucosa made by Bargen in 1935 to the reliable central reading of sigmoidoscopies or multiple colonic segments. Of the endoscopic scoring systems described above, the Mayo Score and UCEIS are reliable and thus currently favored over other scoring systems, but differentiating between these 2 evaluative instruments will require more investigation into their comparative responsiveness and predictive validity. Currently, centrally reading the endoscopic component of the Mayo Score is being used to reduce inclusion bias in clinical trials with the goal of reducing placebo rates, and this strategy seems to have some predictive validity. The UCEIS is often being scored in parallel, given that it is reliable, and has promise as a useful index in the future.

An ideal evaluative instrument for clinical trials will be responsive to effective therapies. Responsive evaluative instruments allow for efficient proof-of-concept studies by facilitating small sample sizes to reach a given power. Furthermore, with responsive evaluative instruments as primary endpoints in UC proof-of-concept studies and the use of statistical techniques such as measuring shifts in distributions of evaluative instrument scores, relatively small sample sizes may have sufficient power to detect a difference between drug and placebo.²⁶

The evidence for the UCCIS discussed above suggests that evaluation of the entire colon with subscores given for each segment might be preferred to assessing the left colon only because it provides a more comprehensive assessment of the colonic mucosa. However, further studies are needed to investigate the feasibility of serial colonoscopy in this clinical setting as well as the added value relative to the current standard of serial flexible sigmoidoscopy. For proof-of-concept studies, an evaluative instrument would ideally minimize cost and risk, which would favor flexible sigmoidoscopic scoring. Nevertheless, the U. S. Food and Drug Administration is evolving toward the requirement of full colonoscopy during phase 3 drug registration if a claim of mucosal healing is desired.

CONCLUSIONS

Both the endoscopic evaluative instrument selected and the definition chosen for mucosal healing affect the validity of assessing endoscopic disease activity during a clinical trial for UC. Currently, the sigmoidoscopic component of the Mayo Score and the UCEIS show the most promise as reliable evaluative instruments of endoscopic disease activity. However, further validation is required.

ACKNOWLEDGMENTS

W. J. Sandborn reports having received consulting fees from Santarus, AbbVie, Actogenix, Boehringer-Ingelheim, Lexicon Pharmaceuticals, Salix Pharmaceuticals, Teva, and Tillotts Pharma and reports having received consulting fees and research grants from Glaxo Smith Kline, Amgen, Bristol Meyers Squibb, Genentech, Hutchison Mediapharma, Janssen, Millennium Pharmaceuticals/Takeda, Pfizer, Prometheus Laboratories, and Receptos. B. G. Feagan reports being a board member of Abbott/AbbVie, Amgen, Astra Zeneca, Avaxia Biologics, Inc, Bristol-Myers Squibb, Celgene, Centocor, Inc, Elan/Biogen, Ferring, JnJ/Janssen, Merck, Novartis, Novonordisk, Pfizer, Prometheus Laboratories, Salix Pharma, Takeda, Teva, Tillotts Pharma AG, and UCB Pharma and reports receiving consulting fees from Abbott/AbbVie, Actogenix, Albireo Pharma, Amgen, Astra Zeneca, Avaxia Biologics, Inc, Axcan, Baxter Healthcare Corp., Boehringer-Ingelheim, Bristol-Myers Squibb, Celgene, Elan/Biogen, EnGene, Ferring Pharma, Roche/Genentech, GiCare Pharma, Gilead, Given Imaging, Inc, GSK, Ironwood Pharma, Janssen Biotech (Centocor), JnJ/Janssen, Kyowa Kakko Kirin Co, Ltd., Lexicon, Lilly, Merck, Millennium, Nektar, Novonordisk, Prometheus Therapeutics and Diagnostics, Pfizer, Receptos, Salix Pharma, Serono, Shire, Sigmoid Pharma, Synergy Pharma, Inc, Takeda, Teva Pharma, Tillotts, UCB Pharma, Warner-Chilcott, Wyeth, Zealand, and Zyngenia. B. G. Feagan also reports receiving research grants from Abbott/AbbVie, Amgen, Astra Zeneca, Bristol-Myers Squibb (BMS), Janssen Biotech (Centocor), JnJ/Janssen, Roche/Genentech, Millennium, Pfizer, Receptos, Santarus, Sanofi, Tillotts, and UCB Pharma. G. R. D'Haens reports having received consulting and/or lecture fees from AbbVie, ActoGeniX, AM Pharma, Boehringer Ingelheim GmbH, Centocor, ChemoCentryx, Cosmo Technologies, Elan Pharmaceuticals, Engene, Dr Falf Pharma, Ferring, Galapagos, Giuliani SpA, Given Imaging, GlaxoSmithKline, Jansen Biologics, Merck Sharp and Dohme Corp, Millennium Pharmaceuticals, Inc (now Takeda), Neovacs, Novonordisk, Otsuka, PDL Biopharma, Pfizer, Receptos, Salix, Setpoint, Shire Pharmaceuticals, Schering-Plough, Tillotts Pharma, UCB Pharma, Versant, and Vifor Pharma and reports receiving research grants from Abbott Laboratories, Jansen Biologics, Given Imaging, MSD, DrFalk Pharma, and Photopill; and speaking honoraria from Abbott Laboratories, Tillotts, Tramedico, Ferring, MSD, UCB, Norgine, and Shire. B. G. Levesque reports having received consulting fees from Prometheus Laboratories and Santarus, Inc. M. A. Samaan, M. H. Mosli, E. Dubcenco, and K. A. Baker have no conflicts of interest to disclose.

Funding was not received from the National Institutes of Health, Wellcome Trust, Howard Hughes Medical Institute, or others.

Authors contributions: M. A. Samaan, M. H. Mosli, B. G. Feagan, W. J. Sandborn, G. R. D'Haens, and B. G. Levesque contributed to the *conception and design of the study, analysis and interpretation of data, and drafting the article*; E. Dubcenco and K. A. Baker contributed to the *analysis and interpretation of the data and revising the manuscript* for important intellectual content. All authors provided final approval of the version to be published.

REFERENCES

- Danese S, Fiocchi C. Ulcerative colitis. *N Engl J Med*. 2011;365:1713–1725.
- Rutgeerts P, Sandborn WJ, Feagan BG, et al. Infliximab for induction and maintenance therapy for ulcerative colitis. *N Engl J Med*. 2005;353:2462–2476.
- Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med*. 2013;369:699–710.
- Froslic KF, Jahnsen J, Moum BA, et al. Mucosal healing in inflammatory bowel disease: results from a Norwegian population-based cohort. *Gastroenterology*. 2007;133:412–422.
- Colombel JF, Rutgeerts P, Reinisch W, et al. Early mucosal healing with infliximab is associated with improved long-term clinical outcomes in ulcerative colitis. *Gastroenterology*. 2011;141:1194–1201.
- Bargen JA. The medical management of chronic ulcerative colitis: (section of surgery: sub-section of proctology). *Proc R Soc Med*. 1937;30:351–362.
- Truelove SC, Witts LJ. Cortisone in ulcerative colitis; final report on a therapeutic trial. *Br Med J*. 1955;2:1041–1048.
- Baron JH, Connell AM, Lennard-Jones JE. Variation between observers in describing mucosal appearances in proctocolitis. *Br Med J*. 1964;1:89–92.
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37–46.
- Sutherland LR, Martin F. 5-Aminosalicylic acid enemas in treatment of distal ulcerative colitis and proctitis in Canada. *Dig Dis Sci*. 1987;32(12 suppl):64S–66S.
- Higgins PD, Schwartz M, Mapili J, et al. Patient defined dichotomous end points for remission and clinical improvement in ulcerative colitis. *Gut*. 2005;54:782–788.
- Schroeder KW, Tremaine WJ, Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. *N Engl J Med*. 1987;317:1625–1629.
- Sandborn WJ, Regula J, Feagan BG, et al. Delayed-release oral mesalazine 4.8 g/day (800-mg tablet) is effective for patients with moderately active ulcerative colitis. *Gastroenterology*. 2009;137:1934–1943. e1–e3.
- Feagan BG, Reinisch W, Rutgeerts P, et al. The effects of infliximab therapy on health-related quality of life in ulcerative colitis patients. *Am J Gastroenterol*. 2007;102:794–802.
- Sandborn WJ, Rutgeerts P, Feagan BG, et al. Colectomy rate comparison after treatment of ulcerative colitis with placebo or infliximab. *Gastroenterology*. 2009;137:1250–1260; quiz 520.
- Levesque B, Pola S, King D, et al. Responsiveness of central endoscopic assessment of disease activity using the Modified Mayo Clinic Score in ulcerative colitis. *Gastroenterology*. 2013;144:S767.
- Travis SP, Schnell D, Krzeski P, et al. Developing an instrument to assess the endoscopic severity of ulcerative colitis: the ulcerative colitis endoscopic index of severity (UCEIS). *Gut*. 2012;61:535–542.
- Travis SP, Schnell D, Krzeski P, et al. Reliability and initial validation of the ulcerative colitis endoscopic index of severity. *Gastroenterology*. 2013;145:987–995.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
- Feagan BG, Sandborn WJ, D'Haens G, et al. The role of centralized reading of endoscopy in a randomized controlled trial of mesalazine for ulcerative colitis. *Gastroenterology*. 2013;145:149–157.
- Bernstein CN, Shanahan F, Anton PA, et al. Patchiness of mucosal inflammation in treated ulcerative colitis: a prospective study. *Gastrointest Endosc*. 1995;42:232–237.
- Kim B, Barnett JL, Kleer CG, et al. Endoscopic and histological patchiness in treated ulcerative colitis. *Am J Gastroenterol*. 1999;94:3258–3262.
- Samuel S, Bruining DH, Loftus EV Jr, et al. Validation of the ulcerative colitis colonoscopic index of severity and its correlation with disease activity measures. *Clin Gastroenterol Hepatol*. 2013;11:49–54.e1.
- Thia KT, Loftus EV Jr, Pardi DS, et al. Measurement of disease activity in ulcerative colitis: interobserver agreement and predictors of severity. *Inflamm Bowel Dis*. 2011;17:1257–1264.
- Su C, Lewis JD, Goldberg B, et al. A meta-analysis of the placebo rates of remission and response in clinical trials of active ulcerative colitis. *Gastroenterology*. 2007;132:516–526.
- Lenth RV. Some practical guidelines for effective sample size determination. *Am Stat*. 2001;55:187–193.

ERRATUM

Crohn's Disease is Associated with Restless Legs Syndrome: ERRATUM

In the article on page 275, volume 16, issue 2, an author's name was listed incorrectly. The second to last author, originally listed as Gerald E. Mullen, MD should appear as: Gerard E. Mullin, MD.

REFERENCE

Weinstock LB, Bosworth BP, Scherl EJ, et al. Crohn's disease is associated with restless legs syndrome. *Inflamm Bowel Dis*. 2010;16:275–279.