



Original Article

Validation of the 'Inflammatory Bowel Disease—Distribution, Chronicity, Activity [IBD-DCA] Score' for Ulcerative Colitis and Crohn's Disease

Corinna Lang-Schwarz,^{a,*} Miriam Angeloni,^{b,*} Abbas Agaimy,^b Raja Atreya,^{c,y} Christoph Becker,^{c,y} Theresa Dregelies,^a Silvio Danese,^d Jean-François Fléjou,^e Nikolaus Gaßler,^f Heike I. Grabsch,^{g,h} Arndt Hartmann,^b Kateřina Kamarádová,ⁱ Anja A. Kühl,^{j,y} Gregory Y. Lauwers,^k Alessandro Lugli,^l Iris Nagtegaal,^m Markus F. Neurath,^{c,y} Georg Oberhuber,^{n,o} Laurent Peyrin-Biroulet,^p Timo Rath,^c Robert Riddell,^q Carlos A. Rubio,^{r,o} Kieran Sheahan,^s Britta Siegmund,^{t,y} Herbert Tilg,^u Vincenzo Villanacci,^v Maria Westerhoff,^w Fulvia Ferrazzi,^{b,x,t} Michael Vieth^{a,b,t}

^aInstitute of Pathology, Klinikum Bayreuth GmbH, Bayreuth, Germany ^bInstitute of Pathology, Friedrich-Alexander-University, Erlangen, Germany ^cDepartment of Medicine & Deutsches Zentrum Immuntherapie DZI, University Hospital, Friedrich-Alexander-University, Erlangen, Germany ^dDepartment of Gastroenterology, Humanitas Research Hospital, Rozzano, Milan, Italy ^ePathology Department, Saint-Antoine Hospital, Pierre et Marie Curie University, Paris, France ^fInstitute for Legal Medicine, Section Pathology, University Hospital, Jena, Germany ^gDepartment of Pathology, School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, The Netherlands ^hPathology and Data Analytics, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, UK ⁱFingerland Department of Pathology, Charles University Faculty of Medicine and University Hospital, Hradec Králové, Czech Republic ^jCharité - Universitätsmedizin Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany ^kDepartment of Pathology, Moffitt Cancer Center, Tampa, FL, USA ^lInstitute of Pathology, University of Bern, Bern, Switzerland ^mDepartment of Pathology, Radboud University Medical Centre, Nijmegen, The Netherlands ⁿInstitute of Pathology, Tirol Kliniken, Innsbruck, Austria ^oInstitute of Pathology, Patho im Zentrum, St. Pölten, Austria ^pDepartment of Gastroenterology, Nancy University Hospital, Lorraine University, Vandoeuvre, France ^qPathology and Laboratory Medicine, Mount Sinai Hospital, University of Toronto, Toronto, ON, Canada ^rDepartment of Pathology, Karolinska Institute and University Hospital, Stockholm, Sweden ^sDepartment of Pathology & Centre for Colorectal Disease, St Vincent's University Hospital and University College, Dublin, Ireland ^tMedical Department [Gastroenterology, Infectiology, Rheumatology], Charité-Universitätsmedizin Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany ^uDepartment of Internal Medicine I, Gastroenterology, Hepatology, Endocrinology and Metabolism, Medical University Innsbruck, Innsbruck, Austria ^vDepartment of Pathology, Spedali Civili Brescia, Brescia, Italy ^wDepartment of Pathology, University of Michigan, Ann Arbor, MI, USA ^xDepartment of Nephropathology, Institute of Pathology, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany ^yTransregio 241 IBDome Consortium, Erlangen, Berlin, Germany

*Corinna Lang-Schwarz and Miriam Angeloni contributed equally to the work and share first authorship.

^tFulvia Ferrazzi and Michael Vieth contributed equally to the work and share last authorship.

Corresponding author: M. Vieth, MD, Institute of Pathology, Klinikum Bayreuth GmbH, Preuschwitzer Str. 101, 95445 Bayreuth, Germany. Tel.: +49 921 400 5602; fax: ++49 921 400 88 5630; email: vieth.lkpathol@uni-bayreuth.de

Abstract

Background and Aims: Histological scoring plays a key role in the assessment of disease activity in ulcerative colitis [UC] and is also important in Crohn's disease [CD]. Currently, there is no common scoring available for UC and CD. We aimed to validate the Inflammatory Bowel Disease [IBD]—Distribution [D], Chronicity [C], Activity [A] score [IBD-DCA score] for histological disease activity assessment in IBD.

Methods: Inter- and intra-rater reliability were assessed by 16 observers on biopsy specimens from 59 patients with UC and 25 patients with CD. Construct validity and responsiveness to treatment were retrospectively evaluated in a second cohort of 30 patients.

Results: Inter-rater reliability was moderate to good for the UC cohort (intraclass correlation coefficients [ICCs] = 0.645, 0.623, 0.767 for D, C, and A, respectively) and at best moderate for the CD cohort [ICC = 0.690, 0.303, 0.733 for D, C, and A, respectively]. Intra-rater agreement ranged from good to excellent in both cohorts. Correlation with the Nancy Histological Index [NHI] was moderate and strong with the Simplified Geboes Score [SGS] and a Visual Analogue Scale [VAS], respectively. Large effect sizes were obtained for all three parameters. External responsiveness analysis revealed correlated changes between IBD-DCA score and NHI, SGS and VAS.

Conclusions: The IBD-DCA score is a simple histological activity score for UC and CD, agreed and validated by a large group of IBD specialists. It provides reliable information on treatment response. Therefore, it has potential value for use in routine diagnostics as well as clinical studies.

Key Words: Histological index; inflammatory bowel disease; IBD-DCA

1. Introduction

Recently, the importance of histological activity scoring in predicting clinical outcomes of patients with inflammatory bowel diseases [IBD] has become apparent. As such, several meta-analyses provide evidence that histological activity scoring can outperform endoscopic activity assessment in ulcerative colitis [UC] in predicting clinical endpoints such as the occurrence of flares, the need for corticosteroid use, and hospitalisation for acute severe UC.^{1–5} In Crohn's disease [CD] with isolated terminal ileum involvement and clinical remission, histological healing was associated with decreased risk of clinical relapse, medication escalation, and need for corticosteroid use.⁶

Furthermore, first evidence suggests that histological activity is superior in predicting the development of dysplasia and carcinoma compared with endoscopic assessment in UC.^{7,8} Histological mucosal healing has emerged as the new ultimate treatment goal in UC and CD, as it is associated with improved clinical outcome, prolonged remission, fewer hospitalisations, and decreased surgery.^{1,2,8–24} However, the most suitable histological target feature to define histological mucosal healing is yet to be found. Recent studies suggest that neutrophils might play a key role in this issue, and other studies focus on architectural distortion as a distinctive feature between histological quiescent disease and true histological normalisation.^{3,25–28}

Therefore, the ideal histological score should be able to assess not only disease activity but also restoration of chronic inflammation to normal, to distinguish between quiescent colitis and histological architectural normalisation. The utility of histological activity scoring is such that, in 2016, the Food and Drug Administration [FDA] of the US Department of Health and Human Services recommended that histological activity scoring in UC should be carried out in parallel with endoscopic assessment.²⁹

Unfortunately there are many IBD scoring systems, some have never been fully validated, others are too complicated for practical use, and many do not fulfill the currently accepted standards for index development.^{30,31} Detailed reviews of as many as 30 histological scoring indices for UC and 13 for CD have been published by the Cochrane Collaboration, highlighting their respective advantages and disadvantages.^{30,31}

Given the urgent clinical need for a usable and standardised histological scoring system in UC and CD, we aimed to develop and validate a simple histological activity scoring index for idiopathic IBD that could be used for clinical trials and routine daily pathology practice and is, at the same time, easy to calculate. In accordance with the existing scores, the new score is not meant to establish a diagnosis of UC or CD but to assess the amount and severity of active and chronic changes in already known IBD.

The Inflammatory Bowel Disease—Distribution, Chronicity, Activity [IBD-DCA] score was initially developed during the International Consensus Conference on Inflammatory Bowel Disease, held in Erlangen, Germany, January 8–10, 2020, with participants from 12 countries.³² The aim of this study was to validate the new score.

2. Methods

The ethics committee of Friedrich-Alexander-University Erlangen-Nuremberg, Germany, approved the study [study number: 175_20 Bc].

2.1. Phase 1: development of the Inflammatory Bowel Disease—Distribution [D], Chronicity [C], Activity [A] [IBD-DCA] score

The IBD-DCA score was proposed during the Consensus Conference. A detailed description regarding how it was developed has already

been published.³² In brief, the score consists of three main parameters which also constitute the name of the new index and the order in which the score should be assessed:

- D for assessment of the distribution of overall active or chronic changes in the IBD colon biopsy, regardless of whether they are epithelial, architectural, or inflammatory;
- C for assessment of features of chronic injury [architectural distortion or chronic inflammation];
- A for assessment of activity features [neutrophils].

For further substratification of the parameters D, C and A, additional items shown to have high inter- or intra-rater reliability in already existing scores in the literature were adopted and included in the new index [Table 1]. An overview of the new IBD-DCA score is shown in Table 2.

A score of D0 implies that C and A are also 0 [normal]. Lymphoid aggregates or lymphoid follicles are part of the normal mucosa and do not qualify for D1 or D2.⁴⁰ If there is only one biopsy, the tissue area of this single biopsy is 100%. If there are more than one biopsy from one container, the tissue area of these biopsies together is 100%.

Scoring should be done for each container separately. However, in case of equal results for all sample sites, the IBD-DCA score might be reported for all biopsies at the end of the report. In this case, it must be stated explicitly that the scoring results were equal for all containers. Detailed recommendations regarding optimal biopsy sampling [including number of biopsies and sites] in order to maximise diagnostic information have also been published by our group.³² Figure 1 shows histological examples for each possible parameter. Figure 2 shows an example for assessment of the IBD-DCA score.

2.2. Phase 2: validation of the IBD-DCA score

2.2.1. Reliability and blinding

To test reliability of using the IBD-DCA score and to validate whether acceptable agreement could be consistently reached among pathologists, digital images of haematoxylin and eosin [H&E] stained biopsies from known UC and CD cases were provided to

16 pathologists. Each pathologist evaluated the biopsies with the IBD-DCA score independently and blinded to patient data as well as to clinical and endoscopic features. The pathologists represented different centres within Europe, the USA, and Canada. Fourteen observers attended the face-to-face meeting, two observers did not. For assessment of intra-observer reliability, eight pathologists scored the entire case series twice with a 3-month washout period in between.

The study set consisted of H&E-stained virtual slides of biopsies from patients with CD [$n = 25$] and UC [$n = 59$] selected randomly from the DC Pathos database [DC Systeme, Heiligenhaus, Germany: <https://www.dc-systeme.de/>] from the Institute of Pathology, Klinikum Bayreuth. The study set represented the full spectrum of disease activity from histological normalisation to severe, ulcerative inflammation. The CD cases included biopsies from colon, terminal ileum, and stomach. All slides had been digitalised at the Institute of Pathology, Klinikum Bayreuth, Germany, using a NanoZoomer S360 scanner [Hamamatsu, Herrsching am Ammersee, Germany]. The participants were given access to the slides online via a password-protected platform using NDPView 2 [Hamamatsu, Herrsching am Ammersee, Germany]. Each pathologist scored every slide for parameters D, C, and A separately, according to Table 2. In addition, participants were asked to record whether an 'A2' was due to crypt abscesses, erosion, or ulceration. These individual items of parameter A2 were finally not included into the IBD-DCA score due to lack of further informative impact.

2.2.2. Feasibility

For assessment of feasibility, a pathologist from a centre in Europe and a pathologist from the USA independently measured the time [in seconds] required to assess the IBD-DCA score for each slide on the virtual scanned slides during the first reading. The measured time includes the time until the slides open and are visible until the scoring is done.

2.2.3. Clinical responsiveness to treatment and construct validity

In addition to the digital slide assessment, the ability of the IBD-DCA score to predict clinical responsiveness to treatment was assessed

Table 1. Histological features used to construct the IBD-DCA score.

Item	ICC [95% CI]	References	Corresponding parameter in IBD-DCA score
Chronic inflammatory infiltrate	0.750 [0.640–1]	NHI ³³	C
	0.75 [0.54–0.86]	RHI ³⁴	
Basal plasmacytosis	0.63 [0.48–0.74]	Mosli <i>et al.</i> ³⁵	C 1 and C 2
Crypt architectural distortion	0.72 [0.59–0.80]	Mosli <i>et al.</i> for MRS and for GS ^{35–37}	C 1
	0.70 [0.56–0.79]		
Acute inflammatory infiltrate	0.772 [0.704–0.940]	NHI ³³	A 1
	0.85 [0.82–0.88]	Bressenot <i>et al.</i> for RI ^{9,38}	
Lamina propria neutrophils	0.61 [0.48–0.69]	Mosli <i>et al.</i> and Bressenot <i>et al.</i> for GS ^{34,37,38}	A 1
	0.82 [0.78–0.86]		
Neutrophils in epithelium	0.74 [0.68–0.80]	Bressenot <i>et al.</i> for GS ^{37,38}	A 1
Erosion	0.79 [0.66–0.86]	RHI ³⁴	A 2
	0.82 [0.77–0.88]	Bressenot <i>et al.</i> for GS ^{37,38}	
Ulceration	0.865 [0.750–1]	NHI ³³	A 2
	0.79 [0.66–0.86]	RHI ³⁴	
	0.82 [0.77–0.88]	Bressenot <i>et al.</i> for GS and for Gramlich Index ^{37–39}	
	0.90 [0.79–0.97]		

IBD-DCA, Inflammatory Bowel Disease—Distribution, Chronicity, Activity; ICC, intraclass correlation coefficient; CI, confidence interval; NHI, Nancy Histological Index; RHI, Roberts Histopathology Index; MRS, Modified Riley Score; GS, Geboes Score; RI, Riley Index.

using cases from a well-characterised cohort of patients by two pathologists [MV and CLS]. Similar to the evaluation performed in the development of the Nancy Histological Index [NHI], responsiveness was retrospectively assessed in 30 patients with UC using two sets

Table 2. Components of the new IBD-DCA score.

Variable	Classification
Distribution [D]	0Normal 1 < 50% of tissue affected per same biopsy site 2 ≥ 50% of tissue affected per same biopsy site
Chronic features [C]	0Normal 1Crypt distortion <i>and/or</i> mild lymphoplasmacytosis 2Marked lymphoplasmacytosis <i>and/or</i> marked basal plasmacytosis
Activity features [A]	0Normal 1Two or more neutrophils in lamina propria in one high-power field [HPF] <i>and/or</i> intraepithelial neutrophils [any number] 2Crypt abscesses, erosions, ulcers

IBD-DCA, Inflammatory Bowel Disease—Distribution, Chronicity, Activity.

of biopsy specimens from each patient, taken at two different time points during treatment. Patients were diagnosed with UC between 2014 and 2020 at the Institute of Pathology at Klinikum Bayreuth GmbH, Bayreuth, Germany. The median time interval between the first and second biopsy was 13 months [3 to 67 months]. Information about treatment was available for 22 patients, most of them receiving combinations of oral and local treatment. Among them, 12 received oral mesalazine [5-ASA] with or without additional topical therapy, three received oral corticosteroids in combination with mesalazine and topical therapy, two were treated with azathioprin, three received tumour necrosis factor alpha [TNF α]-antibody therapy, one was treated with vedolizumab, and one with the JAK-inhibitor tofacitinib.

All patients showed a change in histological disease activity between the two time points [referred to as ‘baseline-condition’ versus ‘follow-up condition’]. Among them:

- $n = 5$ ranged from severe activity [DX, CX, A2] to histological normalisation [D0, C0, A0];
- $n = 15$ ranged from moderate activity [DX, CX, A2] to histological normalisation [D0, C0, A0];
- $n = 5$ ranged from moderate activity [DX, CX, A2] to mild activity [DX, CX, A1]; and
- $n = 5$ ranged from mild activity [DX, CX, A1] to histological normalisation [D0, C0, A0].

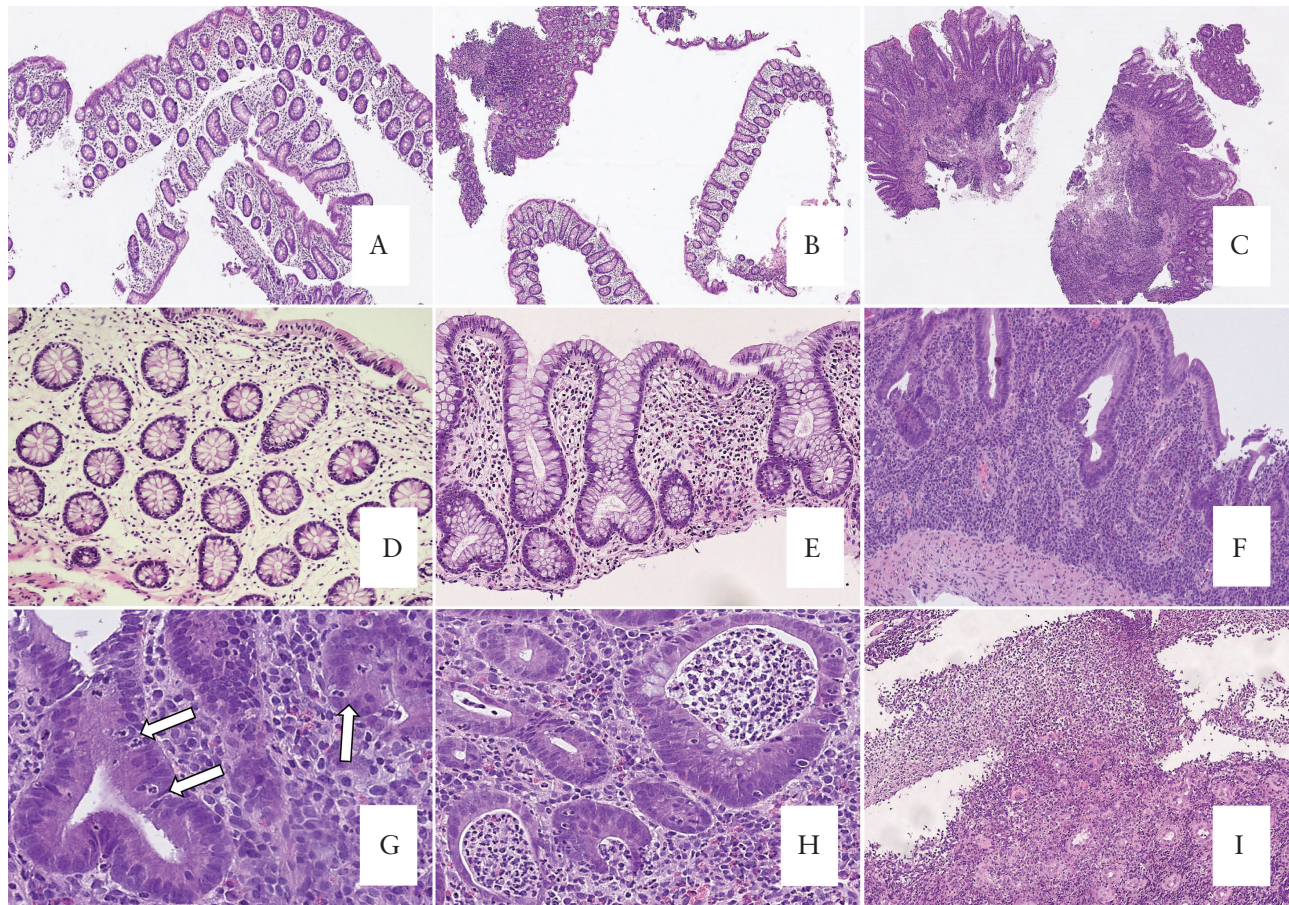
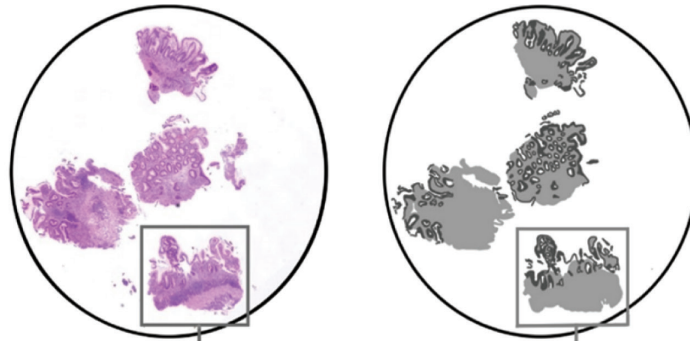


Figure 1. Haematoxylin and eosin [H&E]-stained histological examples for all possible parameters of the IBD-DCA score. a—D0, normal mucosa [and C0 and A0, magnification 8.02x], b—D1, less than 50% of biopsies affected [magnification 3.3x], c—D2, ≥ 50% of biopsies affected [magnification 3.61x], d—C0 and A0, normal mucosa [magnification 20x], e—C1, crypt architectural distortion [magnification 22x], f—C2, architectural distortion and marked lamina propria lymphoplasmacytosis including basal lymphoplasmacytosis [magnification 12.6x], g—A1, intraepithelial neutrophils [white arrows, magnification 35.7x], h—A2, crypt abscesses [h] and ulcer [i] [magnification 33.8x and 16.9x, respectively]. IBD-DCA, Inflammatory Bowel Disease—Distribution, Chronicity, Activity.

1. Distribution: D

Assess parameter D as amount of overall affected tissue in scanning magnification (2.5-4x).

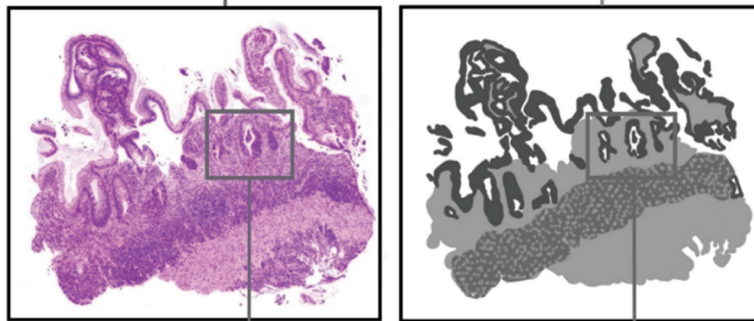
Example shows four biopsies, affected by inflammatory and architectural changes in >50% of tissue, resulting in "D2".



2. Chronicity: C

Assess parameter C in magnification 4 to 10x.

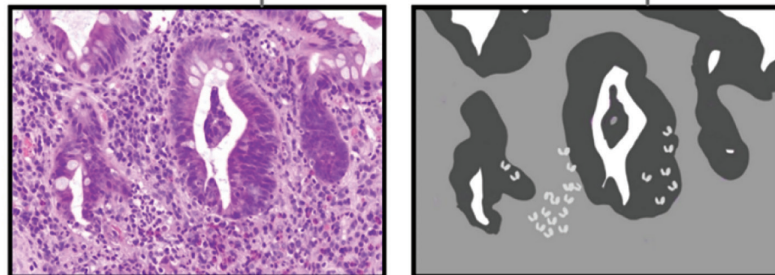
Example shows architectural distortion as well as a particularly prominent bandlike (lympho-) plasmacytosis corresponding to "C2".



3. Activity: A

Assess parameter A in higher magnification.

Example shows a cluster of neutrophilic granulocytes in the tunica propria as well as some granulocytes in the crypt epithelium resulting in "A1".



Summary IBD-DCA score for shown example is: D2 C2 A1.

Figure 2. Example for assessment of the IBD-DCA scoring algorithm with corresponding schemes for better visualisation. IBD-DCA, Inflammatory Bowel Disease—Distribution, Chronicity, Activity.

As the IBD-DCA score summarises severe and moderate activity in its feature A2 due to its two-tiered design, cases with a change in disease activity from severe to moderate were not included. The Mayo Endoscopic Subscore [MES] was also available for each biopsy set. The MES had been assessed by different gastroenterologists with special interest in IBD during routine endoscopy. The slides were retrieved from the archives of the Institute of Pathology, Klinikum Bayreuth, Germany, after a search in the institutional database. A total of 60 slides [two slides per patient] were randomised and the observers were blinded to clinical data and visit number. Each slide was first scored by the two observers using the IBD-DCA score.

In order to compare the IBD-DCA Score with other established scoring systems, the two observers subsequently scored the same

slides using the NHI and the Simplified Geboes Score [SGS], with a minimum of 1 month washout between each scoring.^{33,41} The NHI and the SGS were chosen for comparisons with the IBD-DCA score as the two observers were familiar with those two scores from their participation in clinical trials. In another separate reading, the pathologists again scored the slides evaluating disease activity on a 100-mm Visual Analogue Scale [VAS] with 10 step intervals from 0 [normal mucosa] to 10 [most severe disease activity]. In addition, for each slide, the MES was retrieved from the biopsy submission form for comparison of histological changes in disease activity with the MES.⁴²

In summary, 60 slides of 30 patients were scored four times by the two pathologists in four different readings, with at least 1 month between each reading to exclude a recall bias. As the different indices

IBD-DCA	NHI	SGS	VAS	MES
D0 and C0 and A0	0	0.0	0	0-1
C1 and A0		0.1-1.1	1-2	
C2 and A0	1	1.2	3-4	
	2			
A1	3	2B.1-4.0	5-7	2
A2	4	4.1-4.4	8-10	3

Abbreviations: NHI Nancy Histological Index, SGS Simplified Geboes Score, VAS Visual Analog Scale, MES Mayo Endoscopic Subscore

Figure 3. Conversions of IBD-DCA-Score for correlation analyses versus NHI, SGS, VAS and MES. Correlating scoring values are indicated in same colours. MES grades 0 to 1 both refer to normal mucosa respectively endoscopic remission.

Table 3. Intraclass correlation coefficient estimate [ICC] [with 95% CI] for inter-rater agreement in ulcerative colitis [UC] and Crohn's disease [CD] cohorts after re-scoring.

	ICC [95% CI]	
	UC	CD
D—Distribution	0.645 [0.554-0.737]	0.69 [0.556-0.824]
C—Chronicity	0.623 [0.532-0.717]	0.303 [0.183-0.496]
A—Activity	0.767 [0.695-0.835]	0.733 [0.604-0.852]

CI, confidence interval.

Table 4. Intra-rater agreement for both cohorts.

	UC		CD	
	Median ICC	ICC range	Median ICC	ICC range
D—Distribution	0.894	0.745–1	0.854	0.752–1
C—Chronicity	0.798	0.706–1	0.714	0.442–1
A—Activity	0.909	0.884–0.986	0.888	0.864–1

UC, ulcerative colitis; CD, Crohn's disease; ICC, intraclass coefficient.

cannot be readily converted one to another, conversions between them were established as shown in [Figure 3](#). The construct validity of the IBD-DCA score was evaluated in terms of correlation between the developed score and the NHI, the SGS, and the VAS [as assessed by MV and CLS for responsiveness analyses] as well as the MES.

2.2.4. Statistics

All analyses were performed using the R statistical framework v. 3.6.0.⁴³ The 16 pathologists were given identifiers from 1 to 16. Intra- and inter-rater reliability were assessed separately for items D, C, and A within the two cohorts [UC and CD] by calculating the intraclass correlation coefficient [ICC].⁴⁴ For inter-rater reliability, ICC estimates and their 95% confidence intervals [CI] were calculated using the R package *irr* v.0.84.1 and a single-rating, absolute agreement, two-way random effects model.^{45,46} Analysis for intra-rater reliability was performed using the R package *psych* v.1.9.12.31 and a single-rating, absolute agreement, two-way mixed-effects model.^{46,47} ICC values <0.5, 0.5–0.75, 0.75–0.9, and >0.9 indicated 'poor', 'moderate', 'good', and 'excellent' agreement, respectively.⁴⁵

The construct validity was quantified separately for baseline and follow-up conditions through the pairwise Kendall correlation coefficients [τ_b] between the IBD-DCA score and the other indices

[NHI, SGS, VAS, and MES]. The analysis was performed using the R packages *Kendall* v.2.2 and *NSM3* v.1.14.^{48,49} In order to perform correlation analyses, score values were re-mapped into different multigrade score systems [[Supplementary Tables 1–3, available as Supplementary data at ECCO-JCC online](#)] derived from the correspondence between indices shown in [Figure 3](#).

Construct validity and responsiveness analyses were conducted using the scores from observer MV, after confirming a good inter-rater reliability between CLS and MV for the IBD-DCA score as well as for the other indices [results not shown]. Responsiveness analysis was conducted in terms of internal and external responsiveness. Internal responsiveness evaluates the ability of the IBD-DCA score to predict changes in disease activity between baseline and follow-up condition. Effect size [ES] statistics were used to estimate the magnitude of change. ES was calculated as $ES = Z_w / \sqrt{N}$, where Z_w is the z-score calculated with a Wilcoxon signed rank test for paired samples and N is the number of paired samples, using the R package *coin* v.1.3–1.^{50,51} ES values 0.1–0.3, 0.3–0.5, and >0.5 represent small, moderate, and large changes in the measure, respectively.⁵⁰ External responsiveness evaluates whether changes in the IBD-DCA score correlate with changes in the other scoring systems [NHI, SGS, VAS, and MES]. To this aim, pairwise Kendall correlation between score differences [baseline minus follow-up] was computed.

3. Results

3.1. Reliability

The scoring of the UC cohort showed moderate inter-rater reliability for parameter D [ICC 0.645, 95% CI: 0.554–0.737], poor to moderate agreement for parameter C [ICC 0.568, 95% CI: 0.468–0.673], and moderate to good for parameter A [ICC 0.748, 95% CI: 0.671–0.82]. The scoring of the CD cohort showed an inter-rater agreement from moderate to good for parameters D [ICC 0.655, 95% CI 0.515–0.801] and A [ICC 0.644, 95% CI 0.504–0.792] and poor for parameter C [ICC 0.303, 95% CI 0.183–0.496] [Supplementary Table 4, available as Supplementary data at ECCO-JCC online](#).

Pairwise inter-rater reliability analysis between the 16 raters showed the presence of outliers within the two cohorts, ie, five raters who had poor pairwise agreement with other raters. After having clarified potential misunderstandings of scoring terminology [with consideration of [Table 2](#)], outliers were asked to re-score the parameters [D, C, A] for which they had obtained a poor agreement with other raters. After re-scoring, an improvement in pairwise inter-rater agreement was observed [[Supplementary Figure 1](#) for UC and [Supplementary Figure 2](#) for CD, [available as Supplementary data at ECCO-JCC online](#)] and ICC estimates of the inter-rater reliability across the 16 raters improved [[Table 3](#)].

Table 5. Estimates [with 95% CIs] of pairwise Kendall correlation coefficient [τ_B] between IBD-DCA and other compared indexes in baseline and follow-up conditions.

Compared index	Baseline condition		Follow up condition	
	τ_B [95% CI]	Two-sided <i>p</i> -value	τ_B [95% CI]	Two-sided <i>p</i> -value
NHI	0.595 [0.418–0.773]	<0.001	1 [0.804–1.196]	< 1E-10
SGS	0.792 [0.611–0.972]	2.38E-05	0.963 [0.827–1.1]	< 1E-10
VAS	0.896 [0.721–1.071]	1.19E-06	0.994 [0.859–1.13]	< 1E-10
MES	0.01 [-0.162–0.181]	0.978	0.111 [-0.039–0.26]	0.573

IBD-DCA, Inflammatory Bowel Disease—Distribution, Chronicity, Activity; CI, confidence interval; NHI, Nancy Histological Index; SGS, Simplified Geboes Score; VAS, Visual Analogue Scale; MES, Mayo Endoscopic Subscore.

Table 6. Effect sizes for the histological items of the DCA-score.

	Effect size [ES]	Z-score	<i>p</i> -value
D—Distribution	0.635	3.4765	0.00043
C—Chronicity	1.09	5.972	1.29E-10
A—Activity	1.229	6.731	1.82E-13

DCA, Distribution, Chronicity, Activity.

To assess intra-rater reliability, eight pathologists re-scored the slides a second time. For each parameter, individual intra-rater agreements were summarised by calculating the median [range] ICC. For the UC cohort, an intra-rater agreement of moderate to excellent was reached for parameters D [median ICC 0.894] and C [median ICC 0.798] and of good to excellent for parameter A [median ICC 0.909]. For the CD cohort, an intra-rater agreement of moderate to excellent was reached for parameter D [median ICC 0.854], of poor to excellent for parameter C [median ICC 0.714], and of good to excellent for parameter A [median ICC 0.888]. Intra-rater results for both cohorts are shown in [Table 4](#). Median ICC and range for intra-rater agreement in both cohorts are expressed separately for each histological item [D, C, A].

3.2. Feasibility

The median [range] time required for IBD-DCA assessment for both observers was 20.5 s for the CD cases [7–151 s] and 26.4 s for UC [4.8–300 s].

3.3. Construct validity

Construct validity was evaluated separately for baseline and follow-up conditions relying on THE Kendall correlation coefficient [τ_B]. For the baseline condition, the IBD-DCA score showed a moderate association with the NHI [$\tau_B = 0.595$] and a good association with the SGS [$\tau_B = 0.792$] as well as with the VAS [$\tau_B = 0.896$]. For the follow-up condition, an almost perfect pairwise association of the IBD-DCA score with the NHI [$\tau_B = 1$], the SGS [$\tau_B = 0.963$], and the VAS [$\tau_B = 0.994$] was obtained. In both baseline and follow-up conditions there was no association with the MES [[Table 5](#)].

When comparing the individual MES grades and the IBD-DCA score in terms of number of matches/mismatches after having converted them according to the three-tiered scoring system shown in [Supplementary Table 3](#), it was possible to observe that in the baseline condition, the majority of biopsies was assigned to IBD-DCA grades 1 and 2 irrespectively of their MES grade, whereas in follow-up condition almost all biopsies were assigned to IBD-DCA grade 0 in accordance with the MES. Correlations between

the individual MES grades with the IBD-DCA score are available as [Supplementary Table 5](#), available as [Supplementary data at ECCO-JCC online](#).

3.4. Responsiveness

In internal responsiveness analysis, all three histological parameters [D, C, and A] showed a large magnitude of change [ES = 0.635, 1.09, 1.229 for D, C, and A, respectively] [[Table 6](#)].

Changes in histopathological scores showed good degrees of correlation with each other, whereas correlation between histopathological scores and Mayo endoscopic subscore was much lower [[Table 7](#)].

4. Discussion

The number of existing scoring systems for assessment of histological activity in ulcerative colitis and Crohn's disease seems—at first glance—large enough to provide ‘the perfect index’ to every pathologist and corresponding clinician. To date, pathologists are either free in their choice of index and even whether to use them at all or make this choice jointly with their endoscopists.^{33–35,52}

Nevertheless, when considered more closely, the existing indices show limitations. The main limitation concerning the indices for CD is that none of them has been fully validated to date.³¹ Concerning UC, existing indices are very heterogeneous in their complexity of assessment algorithms as well as their content of assessed items.^{30,34} The Nancy Histological Index [NHI] and the Robarts Histopathology Index [RHI] have undergone the most validation for UC so far, but both include inflammatory features only, as architectural features were thought unlikely to be responsive to change following therapy.^{31,33}

Histological mucosal healing is not well defined to date. In their currently published position paper, the European Crohn's and Colitis Organisation define histological remission in its strictest way as return to normal.⁵³ Therefore, crypt architectural distortion might be one of the new key features in this issue, as it distinguishes between quiescent UC [which has architectural distortion] and true histological normalization [which looks like normal colon].^{25,26}

This is especially strengthened by the findings of Christensen *et al.*, who showed increased odds of relapse-free survival for histological normalisation in comparison with endoscopic healing or histological quiescence, in a large cohort of 646 patients.³ In a recently published systematic review and meta-analysis including 28 studies with 2806 patients [2677 with UC and 129 with CD], crypt architectural irregularities were also one of the individual features that predicted relapse, as were basal plasmacytosis, neutrophilic infiltrations, and mucin depletion.⁵ Concerning CD, the role of histology in activity assessment is not definitely clear yet. However,

Table 7. Kendall correlation coefficient [95% CI] between score changes in histopathological and endoscopic scores.

	NHI	SGS	VAS	MES
DCA	0.753 [0.574 to 0.933]	0.89 [0.722 to 1.058]	0.953 [0.813 to 1.093]	0.11 [-0.102 to 0.322]
NHI		0.565 [0.338 to 0.792]	0.739 [0.571 to 0.908]	0.288 [0.064 to 0.513]
SGS			0.841 [0.671 to 1.011]	0.1 [-0.121 to 0.321]
VAS				0.27 [0.075 to 0.465]

DCA, Distribution, Chronicity, Activity; CI, confidence interval; NHI, Nancy Histological Index; SGS, Simplified Geboes Score; VAS, Visual Analogue Scale; MES, Mayo Endoscopic Subscore.

according to Christensen *et al.*, histological healing has also shown to be superior in predicting clinical outcomes in ileal CD in comparison with endoscopic healing.⁶

As activity in UC patients should be assessed by endoscopy in conjunction with histology, according to current FDA recommendations, the participants of the International Consensus Conference on Activity Scoring in IBD developed a histological score for both UC and CD which could easily be implemented in routine daily practice and is able to distinguish between histological remission and histological normalisation.

A major strength of the proposed new IBD-DCA score lies in its good inter- and intra-rater reliability, which was assessed in—to the best of our knowledge—the largest group of pathologists to date. Without special training in scoring, inter-rater reliability revealed moderate ICC estimates for all three parameters (D [distribution], C [chronicity] and A [activity]) within the UC cohort, and moderate ICC estimates for parameters D and A within the CD cohort. These values further improved when four out of the five initial ‘outlier’ raters performed a second round of scoring. One of these ‘outliers’ did not attend the face-to-face meeting, which could partially explain the initial poor agreement with the other raters. However, it is worth noting that the results of observer number 11, who also did not attend the meeting, were consistent with those of the other raters from the very beginning, emphasising the simplicity of the score. Overall, the IBD-DCA score reached comparable ICC estimates to those of other published indices, despite including four to five times more observers.^{33–37,54}

To the best of our knowledge and according to Mosli *et al.*, none of the established scoring indices assessed feasibility so far, nor did the latest developed Robarts Histopathology Index [RHI].^{30,34} In this study, feasibility assessment was done on whole slide images [virtual slides], demonstrating that the IBD-DCA score is applicable in digital pathology. Scoring of virtual slides is a pillar of digital pathology and will be probably used more frequently in the future.

Another advantage of the IBD-DCA score, compared with other scoring systems, is its simplicity. It is configured to intuitively follow the usual practice of a pathologist assessing a bowel biopsy specimen from low to high power magnification. Features of chronic injury as well as active inflammatory findings are—apart from normal—only divided into two levels of severity, creating a two-tiered system for chronic as well as active inflammation. The advantage of a two-tiered system is strengthened by the findings of Lemmens *et al.*⁵⁵ In their correlation analysis between endoscopic and histological scores, some scores were infrequently used. This was especially true for the middle grades. The fact that inter-rater agreement in the IBD-DCA score for parameter A was good for UC, but moderate for the stratification of A2 into its special features, further confirms this proposed concept. Erosions and ulcers have also been summarised as ‘mucosal breaks’ in other gastrointestinal diseases, due to lack of negative clinical impact.⁵⁶

Our study had some limitations. The main limitation is that the validation was performed retrospectively on slides from routine work. Further prospective validation in additional datasets, preferably from a randomised controlled trial, is clearly necessary. Although the CD cohort data are promising, there is also clearly a need for further prospective validation in larger cohorts for the upper and lower gastrointestinal tract, including a study set for responsibility analysis to prove potential applicability of the IBD-DCA score for CD, as the role of histological activity assessment in CD is not yet definitely clear due to the discontinuous and transmural nature of the disease.

In this study, we presented the IBD-DCA score that has been developed with international consensus and validated in its interobserver agreement by a large group of pathologists from Europe, the USA, and Canada. Although further studies are clearly necessary, our findings open new avenues for the clinical use of the IBD-DCA score for routine use in histological assessment of IBD activity.

Additional data, as far as not published, are available on demand via email, from the corresponding author

Funding

This work was supported by the Deutsche Forschungsgemeinschaft TRR241 projects INF, A03, C02, C03, and C04 [grant number 375876048]. The Deutsche Forschungsgemeinschaft funds the Heisenberg Professorship of RA [grant number 316866639].

Conflict of Interest

MV reports lecture fees from Falk, Shire, Lilly, Malesci, Pentax, Olympus, and AstraZeneca. RA has served as a speaker or consultant or received research grants from AbbVie, Biogen, Boehringer Ingelheim, Celgene, Celltrion, Dr Falk Pharma, Ferring, InDex Pharmaceuticals, Janssen-Cilag, MSD Sharp & Dome, Pfizer, Roche Pharma, Samsung Bioepis, Takeda. LP-B reports personal fees from Galapagos, AbbVie, Janssen, Genentech, Ferring, Tillots, Pharmacosmos, Celltrion, Takeda, Boehringer Ingelheim, Pfizer, Index Pharmaceuticals, Sandoz, Celgene, Biogen, Samsung Bioepis, Alma, Sterna, Nestle, Inotrem, Enterome, Allergan, MSD, Roche, Arena, Gilead, Hikma, Amgen, BMS, Vifor, Norgine; Mylan, Lilly, Fresenius Kabi, Opillan Pharma, Sublimity Therapeutics, Applied Molecular Transport, OSE Immunotherapeutics, Entera, Theravance; grants from Abbvie, MSD, Takeda; stock options: CTMA. GYL is a consultant for Alimentiv. BS has served as consultant for Abbvie, BMS, Boehringer, Celgene, Falk, Janssen, Lilly, Pfizer, Prometheus, Takeda; and has received speaker’s fees from Abbvie, CED Service GmbH, Falk, Ferring, Janssen, Novartis, Pfizer, and Takeda as a representative of Charité – Universitätsmedizin Berlin.

Author Contributions

Concept and design of the study: CL-S, MV, MFN, AH. Acquisition of data: CL-S, AA, J-FE, HIG, AH, KK, AAK, GYL, AL, IN, GO, RR, CAR, KS, VV, MW, MV. Article drafting: CL-S, MW. Analysis and interpretation of data: MA, FE. Critical revision of the article for important intellectual content: RA,

CB, TD, SD, MFN, LP-B, TR, RR, BS, HT, MW. Final approval of the submitted version: all authors.

Acknowledgements

The participants thank Tilman Schulz, MD, Institute of Pathology, Klinikum Bayreuth GmbH, Bayreuth, Germany, for providing us with the schemes in Figure 2.

Supplementary Data

Supplementary data are available at ECCO-JCC online.

References

- Park S, Abdi T, Gentry M, Laine L. Histological disease activity as a predictor of clinical relapse among patients with ulcerative colitis: systematic review and meta-analysis. *Am J Gastroenterol* 2016;111:1692–701.
- 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.
- Christensen B, Hanauer SB, Erlich J, et al. Histological normalization occurs in ulcerative colitis and is associated with improved clinical outcomes. *Clin Gastroenterol Hepatol* 2017;15:1557–64.e1.
- Yoon H, Jangi S, Dulai PS, et al. Incremental benefit of achieving endoscopic and histological remission in patients with ulcerative colitis: a systematic review and meta-analysis. *Gastroenterology* 2020;159:1262–75.e7.
- Gupta A, Yu A, Peyrin-Biroulet L, Ananthakrishnan AN. Treat to target: the role of histological healing in inflammatory bowel diseases: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2020;30:S1542–3565[20]31376–8.
- Christensen B, Erlich J, Gibson PR, Turner JR, Hart J, Rubin DT. Histological healing is more strongly associated with clinical outcomes in ileal Crohn's disease than endoscopic healing. *Clin Gastroenterol Hepatol* 2020;18:2518–25.e1.
- Rubin DT, Huo D, Kinnucan JA, et al. Inflammation is an independent risk factor for colonic neoplasia in patients with ulcerative colitis: a case-control study. *Clin Gastroenterol Hepatol* 2013;11:1601–8.e1.
- Gupta RB, Harpaz N, Itzkowitz S, et al. Histological inflammation is a risk factor for progression to colorectal neoplasia in ulcerative colitis: a cohort study. *Gastroenterology* 2007;133:1099–105; quiz 1340–1.
- Riley SA, Mani V, Goodman MJ, Dutt S, Herd ME. Microscopic activity in ulcerative colitis: what does it mean? *Gut* 1991;32:174–8.
- Bitton A, Peppercorn MA, Antonioli DA, et al. Clinical, biological, and histological parameters as predictors of relapse in ulcerative colitis. *Gastroenterology* 2001;120:13–20.
- Neurath MF, Travis SP. Mucosal healing in inflammatory bowel diseases: a systematic review. *Gut* 2012;61:1619–35.
- Truelove SC, Richards WC. Biopsy studies in ulcerative colitis. *Br Med J* 1956;1:1315–8.
- Frøslie KF, Jahnsen J, Moum BA, Vatn MH; IBSEN Group. Mucosal healing in inflammatory bowel disease: results from a Norwegian population-based cohort. *Gastroenterology* 2007;133:412–22.
- 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;141:1194–201.
- Sandborn WJ, Rutgeerts P, Feagan BG, et al. Colectomy rate comparison after treatment of ulcerative colitis with placebo or infliximab. *Gastroenterology* 2009;137:1250–60.
- 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.
- 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;110:1324–38.
- Bessissow T, Lemmens B, Ferrante M, et al. Prognostic value of serologic and histologic markers on clinical relapse in ulcerative colitis patients with mucosal healing. *Am J Gastroenterol* 2012;107:1684–92.
- Hefti MM, Chessin DB, Harpaz NH, Steinhagen RM, Ullman TA. Severity of inflammation as a predictor of colectomy in patients with chronic ulcerative colitis. *Dis Colon Rectum* 2009;52:193–7.
- Peyrin-Biroulet L, Bressenot A, Kampman W. Histologic remission: the ultimate therapeutic goal in ulcerative colitis? *Clin Gastroenterol Hepatol* 2014;12:929–34.
- Feagins LA, Melton SD, Iqbal R, Dunbar KB, Spechler SJ. Clinical implications of histologic abnormalities in colonic biopsy specimens from patients with ulcerative colitis in clinical remission. *Inflamm Bowel Dis* 2013;19:1477–82.
- Baert F, Moortgat L, Van Assche G, et al.; Belgian Inflammatory Bowel Disease Research Group; North-Holland Gut Club. Mucosal healing predicts sustained clinical remission in patients with early-stage Crohn's disease. *Gastroenterology* 2010;138:463–8; quiz e10–1.
- De Cruz P, Kamm MA, Prideaux L, Allen PB, Moore G. Mucosal healing in Crohn's disease: a systematic review. *Inflamm Bowel Dis* 2013;19:429–44.
- Schnitzler F, Fidder H, Ferrante M, et al. Mucosal healing predicts long-term outcome of maintenance therapy with infliximab in Crohn's disease. *Inflamm Bowel Dis* 2009;15:1295–301.
- Geboes K, Dalle I. Influence of treatment on morphological features of mucosal inflammation. *Gut* 2002;50[Suppl 3]:III37–42.
- Schumacher G, Kollberg B, Sandstedt B. A prospective study of first attacks of inflammatory bowel disease and infectious colitis. Histological course during the 1st year after presentation. *Scand J Gastroenterol* 1994;29:318–32.
- 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;18:2510–17.e5.
- Magro F, Esteve MM. Do neutrophils contribute to development of Crohn's disease and ulcerative colitis? *Clin Gastroenterol Hepatol* 2020;18:2430–1.
- Food and Drug Administration. *Ulcerative Colitis: Clinical Trial Endpoints Guidance for Industry. Draft guidance*. Silver Spring, MD: FDA; 2016; p19.
- Mosli MH, Parker CE, Nelson SA, et al. Histological scoring indices for evaluation of disease activity in ulcerative colitis. *Cochrane Database Syst Rev* 2017;5:CD011256.
- Novak G, Parker CE, Pai RK, et al. Histological scoring indices for evaluation of disease activity in Crohn's disease. *Cochrane Database Syst Rev* 2017;7:CD012351.
- Lang-Schwarz C, Agaimy A, Atreya R, et al. Maximizing the diagnostic information from biopsies in chronic inflammatory bowel diseases: recommendations from the Erlangen International Consensus Conference on Inflammatory Bowel Diseases and presentation of the IBD-DCA score as a proposal for a new index for histological activity assessment in ulcerative colitis and Crohn's disease. *Virchows Arch* 2021;478:581–94.
- Marchal-Bressenot A, Salleron J, Boulagnon-Rombi C, et al. Development and validation of the Nancy histological index for UC. *Gut* 2017;66:43–9.
- Mosli MH, Feagan BG, Zou G, et al. Development and validation of a histological index for UC. *Gut* 2017;66:50–8.
- Mosli MH, Feagan BG, Zou G, et al. Reproducibility of histological assessment of disease activity in UC. *Gut* 2014;1–9. doi:10.1136/gutjnl-2014-307536.
- Feagan BG, Greenberg GR, Wild G, et al. Treatment of ulcerative colitis with a humanized antibody to the alpha4beta7 integrin. *N Engl J Med* 2005;352:2499–507.
- Geboes K, Riddell R, Ost A, Jensfelt B, Persson T, Löfberg R. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. *Gut* 2000;47:404–9.
- Bressenot A, Salleron J, Bastien C, Danese S, Boulagnon-Rombi C, Peyrin-Biroulet L. Comparing histological activity indexes in UC. *Gut* 2015;64:1412–8.
- Gramlich T, Petras RE. Pathology of inflammatory bowel disease. *Semin Pediatr Surg* 2007;16:154–63.

40. Levine DS, Haggitt RC. Normal histology of the colon. *Am J Surg Pathol* 1989;13:966–84.
41. Jauregui-Amezaga A, Geerits A, Das Y, et al. A simplified Geboes score for ulcerative colitis. *J Crohns Colitis* 2017;11:305–13.
42. Rutgeerts P, Sandborn WJ, Feagan BG, et al. Infliximab for induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2005;353:2462–76.
43. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing; 2019.
44. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
45. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155–63.
46. Gamer M, Lemon J, Gamer MM, Robinson A. Kendall's, "Package 'irr'," *Various coefficients of interrater reliability and agreement*. 2012.
47. Revelle WR. "psych: Procedures for personality and psychological research". 2017.
48. McLeod A, McLeod MA. "Package 'Kendall'". 2015.
49. Schneider G, Chicken E, Becvarik R, Schneider MG. "Package 'NSM3'". 2020.
50. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen* 2012;141:2–18.
51. Hothorn T, Hornik K, Wiel M, van de Zeileis A. Implementing a class of permutation tests: the coin package. *J Stat Softw* 2008;28.
52. Coolican H. *Research Methods and Statistics in Psychology*. London: Psychology Press; 2017.
53. Magro F, Doherty G, Peyrin-Biroulet L, et al. ECCO position paper: harmonization of the approach to ulcerative colitis histopathology. *J Crohns Colitis* 2020;14:1503–11.
54. 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;68:1162–8.
55. Lemmens B, Arijis I, Van Assche G, et al. Correlation between the endoscopic and histologic score in assessing the activity of ulcerative colitis. *Inflamm Bowel Dis* 2013;19:1194–201.
56. Armstrong D, Bennett JR, Blum AL, et al. The endoscopic assessment of esophagitis: a progress report on observer agreement. *Gastroenterology* 1996;111:85–92.