Radboud Repository

Radboud University Nijmegen

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a postprint version which may differ from the publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/191267

Please be advised that this information was generated on 2019-11-22 and may be subject to change.

Journal of Crohn's and Colitis, 2017, 1–7
doi:10.1093/ecco-jcc/jjx165
Original Article

Original Article

Assessment of Histological Remission in Ulcerative Colitis: Discrepancies Between Daily Practice and Expert Opinion

Tessa E. H. Römkens[a,b], Pim Kranenburg[a], Arjan van Tilburg[c], Carolien Bronkhorst[d], Iris D. Nagtegaal[c], Joost P. H. Drenth[a], Frank Hoentjen[a]

[a]Department of Gastroenterology and Hepatology, Radboud University Medical Centre, Nijmegen, The Netherlands [b]Department of Gastroenterology and Hepatology, Jeroen Bosch Hospital, 's Hertogenbosch, The Netherlands [c]Department of Pathology, Radboud University Medical Center, Nijmegen, The Netherlands [d]Department of Pathology, Jeroen Bosch Hospital, 's Hertogenbosch, The Netherlands

Corresponding author: Tessa E. H. Römkens, Radboud University Nijmegen Medical Centre, Department of Gastroenterology and Hepatology, PO Box 9101, 6500 HB Nijmegen, the Netherlands. Tel: +31 243614760; fax +31 243540103 email: tessa.romkens@radboudumc.nl

Conference presentation: Part of this work appeared as an oral presentation at the Digestive Disease days, Veldhoven, the Netherlands, October 2017, and as a poster presentation at the 25th United European Gastroenterology [UEG] week, Barcelona, Spain, October 2017.

Abstract

Background and Aims: Histological remission [HR] is a potential treatment target in ulcerative colitis [UC]. Limited 'real world' data are available on the reliability of histological scoring when assessing minimal histological inflammation. The aim of this study was to investigate the reliability of UC histological scores in colonic biopsies showing mucosal healing [MH] and limited histological inflammation, and to compare the 'daily practice' histological assessment with expert reviews by gastrointestinal [GI] pathologists.

Methods: We performed a retrospective single-centre study. Colonic biopsies from UC patients with MH [Mayo score ≤ 1] were included. All biopsies assessed in daily practice were reassessed by three blinded GI pathologists using three histological scores (Geboes score [GS], Riley score [RS], Harpaz [Gupta] Index [HGI]) and a global visual scale [GVS]. We evaluated inter- and intra-observer variation between GI pathologists and correlations between scores including the initial histological assessment using Cronbach's alpha and Spearman rho analysis.

Results: In total, 270 biopsies from 39 UC patients were included. The inter-observer concordance for all histological indexes was substantial to almost perfect [GS 0.84; HGI 0.61; GVS 0.74, RS 0.91]. Correlation between the RS and GS was almost perfect [R = 0.86], but we found no correlation between the primary histological assessment and reassessment by GI pathologists.

Conclusions: Current UC histological scores reliably assess limited histological inflammation in UC patients. The discrepancy between the initial histological assessment and the reassessment by dedicated GI pathologists suggests a gap between daily practice and academic expertise. This issue may limit the implementation of HR as a treatment target for UC in daily practice.

Key Words: Histological remission; ulcerative colitis; pathology; central reading

Copyright © 2017 European Crohn's and Colitis Organisation (ECCO). Published by Oxford University Press. All rights reserved.
For permissions, please email: journals.permissions@oup.com

Downloaded from https://academic.oup.com/ecco-jcc/advance-article-abstract/doi/10.1093/ecco-jcc/jjx165/4733426
by University of Florida user
on 05 January 2018

1. Introduction

Treatment goals for inflammatory bowel disease [IBD] have changed from assessing primarily symptom-based treatment to evaluating treatment targets that combine both patient-reported outcome [PRO] remission and endoscopic remission (mucosal healing [MH]).[1] The advent of biologicals with positive effects on mucosal inflammation has contributed significantly to this shift.[2–6] MH is a well-accepted treatment target in clinical IBD trials[7–10] and is associated with relevant clinical outcomes.[6,11–13] Some clinical trials have introduced even more stringent treatment targets in IBD such as 'histological remission [HR]'. HR might be a better predictor of a more favourable disease course, in particular in ulcerative colitis [UC].[14–18] However, there is no unified definition of HR,[15,19,20] which seems pivotal for the design and comparability of clinical trials, but also for the use of HR as a treatment target in daily practice. Several scoring indexes are available to measure histological activity in UC. Frequently used scores include the Geboes score [GS], Riley score [RS] and to a lesser extent the easy to use Harpaz [Gupta] Index [HGI]. Only limited data on the reproducibility of these scores were available until recently.[21–23] Two studies evaluating the reproducibility and reliability of these histological activity indexes showed a strong correlation between scores and good intra-observer reproducibility with moderate to good inter-observer agreement.[24,25] The greater portion of patients in both studies, however, had [mildly] active disease. Indeed, the histological item that showed the strongest intra- and inter-observer agreement was 'erosion/ulceration', a feature that correlates with active disease. In addition, the assessing pathologists were all expert gastrointestinal [GI] pathologists. This leads us to two questions. First, does this high observer agreement still hold when re-evaluating colonic biopsies with MH that were initially assessed by a general pathologist as 'HR'. Secondly, do these histological indexes, so far mainly used in clinical trials, correlate with the histological assessment in daily practice by a general pathologist? To assess these questions, we designed a study to compare the reproducibility and reliability of three histological scoring indexes [GS, RS, HGI] and one global visual score [GVS] through intra- and inter-observer agreement testing, in colonic biopsies with MH that were primarily assessed as HR in daily practice. We also investigated the correlation between histological assessment of UC biopsies in daily practice and the re-assessment by specialized GI pathologists.

2. Materials and Methods

2.1. Study design

We designed a retrospective single-centre study in a tertiary referral centre. This trial [ISRCTN61139227] was approved by the Ethics Committee of Radboud University Medical Centre, Nijmegen, the Netherlands.

2.2. Patients

We searched our endoscopy database to identify patients with established UC who met the inclusion criteria. Patients were included if they had [i] a colonoscopic examination between January 2014 and July 2015, [ii] a well-established diagnosis of UC according to clinical and histological criteria [Montreal classification E1–3 were included[26]], [iii] endoscopic MH throughout the examined colon according to the Mayo endoscopic score ≤1[13,27,28] – after inclusion, two groups were distinguished based on the endoscopic Mayo score [0 vs 1] and [iv] obtained and well-documented colonic biopsies from rectum, sigmoid and proximal colon [proximal of splenic flexure]. Patients were excluded if they had [i] a Mayo endoscopic score >1 or [ii] a diagnosis of either Crohn's disease or IBD unclassified. Demographic data of patients were anonymously collected from digital patient records.

2.3. Assessment of biopsy specimens

Biopsies were paraffin-embedded, sectioned and H&E-stained. The slides were scanned at 200× magnification. All biopsies were reassessed by two blinded expert GI pathologists [IN, CB] and one GI pathology fellow [AT]. Two pathologists are employed in an academic hospital [IN, AT] and one in a large non-academic teaching hospital [CB]. Prior to the assessment a consensus meeting was held and teaching materials with sample images of the different scores were developed. Three histological scoring systems [GS, RS and HGI] were used during reassessment, as well as a GVS, indicating a 'first glance' assessment. The GVS is a visual scale ranging from 0 [no activity] to 10 [maximal activity]. All three scoring systems are described in detail in Supplementary Data Tables S1–S3. The GS[21] assesses five features which result in a score ranging from 0 to 5.4, with higher scores indicating more inflammation. The original RS evaluates six features which are scored on a four-point scale [none, mild, moderate or severe], in which equal weight to all six measures was given.[22] The HGI[23] is based on three features resulting in a four-point scale [Histological Activity Index 0 to 3] with higher scores indicating greater inflammation.

Each pathologist scored all biopsies by sections that represent a specific part of the colon [rectum, left-sided or proximal of splenic flexure]. Each section contained several biopsies, with the worst score applied for further analysis. Two blinded pathologists [IN, AT] scored 18 random sections twice, with an interval of at least 2 weeks, to evaluate intra-observer reliability. From all biopsies obtained during colonoscopy, the report of the 'general pathologist [gp]' [the general pathologist who initially assessed the biopsies in daily clinical setting] was evaluated and used as a baseline to construct two groups for analysis. Biopsies that were initially evaluated as 'histological remission' [gpHR] were compared with those showing 'histological inflammation' [gpHI], with regard to inter-observer agreement and intra-observer reliability for all scores. In addition, we assessed the correlation between the primary histological assessment and the different histological scores as well as the correlation between endoscopic Mayo score and histological scores. Biopsies were divided and compared by location [rectum, left-sided colon, colon proximal of splenic flexure] and Montreal classification [E1–3].

2.4. Statistical analysis

SPSS version 22.0 and Scilab were used for the statistical analysis. We used descriptive statistics to analyse the results using counts and proportions for categorical data and means and standard deviations for continuous variables. We evaluated inter- and intra-observer variation of GI pathology experts for GS, HGI and GVS total scores in biopsies taken from endoscopically healed mucosa with Cronbach's alpha statistics, and for the 95% confidence interval we used bootstrapping and provided the percentage of agreement. We used the interpretation of kappa values as suggested by Landis[29]: 0.00 poor; 0.00–0.20 slight; 0.21–0.40 fair; 0.41–0.60 moderate; 0.61–0.80 substantial; 0.81–1.00 almost perfect. For the RS we evaluated inter- and intra-observer concordance between the six different items of the RS with Cronbach's alpha statistics. Where there was almost perfect inter-observer concordance, we provide the percentage of agreement, because Cronbach's alpha statistics could not reliable be used. The relationship between indexes was studied by Spearman's correlation coefficient. We compared the outcome of biopsies that are primarily assessed as 'normal' [gpHR] or 'abnormal' [gpHI] by the general pathologist, and GI pathology experts.

Downloaded from https://academic.oup.com/ecco-jcc/advance-article-abstract/doi/10.1093/ecco-jcc/jjx165/4733426
by University of Florida user
on 05 January 2018

3. Results

3.1. General

We assessed 77 sections with 270 H&E-stained biopsies (143 rectum [53%], 20 sigmoid [7%], 107 proximal of splenic flexure [40%]) from 43 endoscopies [39 colonoscopies, four sigmoidoscopies] in 39 patients. The median number of biopsies was 4.61 in the rectum and 3.34 in the proximal colon. Sixty-three sections [82%] were initially assessed by the general pathologist as histologically healed [gpHR]. Demographic variables of the included patients and details on endoscopy are summarized in Table 1.

3.2. Endoscopic Mayo score – histological activity scores

In 82% [32/39] of the colonoscopies we found an endoscopy score of Mayo 0 in the entire colon, with 100% of patients [39/39] showing Mayo 0 in the colon proximal of the splenic flexure. In 33/43 endoscopies we found a Mayo score of 0 in the rectosigmoid [76.7%]. Eighty-seven per cent [67/77] of all evaluated sections were biopsy sections taken from areas in the colon classified as endoscopy score Mayo 0, with 13% [10/77] scored as Mayo 1.

3.3. Endoscopy and histology scores – general pathologist vs expert GI-pathologists

The general pathologist assessed 84.8% of the rectosigmoid biopsies, endoscopically scored as Mayo 0, as gpHR; in the proximal colon biopsies this was 94.1% [Table 2]. For endoscopy score Mayo 1, the general pathologist assessed 20% [2/10] of the biopsies in the rectum and sigmoid as gpHR.

The results of the re-evaluation of the Mayo 0 and 1 biopsies by expert GI-pathologists using three histological scores [GS, HGI, RS] are depicted in Table 2. To calculate these percentages for all three GI-pathologists, we used the sum of all biopsies. The used cut-offs of the GS [<0.1, <3.1] were used in concordance with the literature.[15]

Table 1. Demographic variables and details on endoscopy.

Patient characteristics [*N* = 39]		
Age [years] [mean ± SD]	53.4 ± 16.09	
Duration of UC [years] [mean ± SD]	23.0 ± 12.9	
Male gender [*N*, %]	23 [59%]	
Montreal classification [*N*, %]	E1	6 [15.4%]
	E2	12 [30.8%]
	E3	21 [53.8%]
Medication [*N*, %]	Mesalamine	28 [71.8 %]
	Thiopurines	13 [33.3%]
	Corticosteroids	2 [5.1%]
	Anti-TNF*	3 [7.7%]
	Topical therapy	7 [17.9%]
Endoscopy [*N* = 43]		
Sigmoidoscopy [*N*, %]	4 [9.3]	
Colonoscopy [*N*, %]	39 [90.7]	
Mayo score 1 rectum	9 [20.1]	
Mayo score 1 rectosigmoid	10 [23.3%]	
Mayo score 1 proximal of splenic flexure††	0 [0%]	

*One patient used combination thiopurine/infliximab; ††*N* = 39 colonoscopies.

3.4. Inter-observer agreement expert GI-pathologists

For all evaluated histological scores, inter-observer concordance was higher in sections assessed as Mayo score 1 vs sections assessed as Mayo score 0 [Table 3 and Table S4].

3.5. Harpaz [Gupta] index [HGI]

Overall inter-observer concordance in the HGI was substantial [0.62]. In primary assessment both as gpHR [0.68] and as gpHI [0.63] this was substantial. When the evaluated sections were divided according to Mayo score, the inter-observer concordance for HGI was moderate for Mayo score 0 [0.55] and substantial for Mayo score 1 [0.64] [Table 3].

3.6. Geboes [GS]

Overall inter-observer concordance for the GS with and without subscores was almost perfect [0.84 and 0.83]. When the primary histological assessment consisted of gpHR the inter-observer concordance was substantial [0.65 and 0.63], and for gpHI there was a higher concordance [0.91 and 0.91]. When the evaluated sections were distinguished according to the endoscopic Mayo score, the inter-observer concordance for GS with subscoring was substantial for Mayo score 0 [0.7] and almost perfect for Mayo score 1 [0.93], and for GS without subscoring the result was similar [0.68 and 0.92] [Table 3].

3.7. Riley score [RS]

Inter-observer concordance was determined for each separate item of the RS. The inter-observer concordance for all six items was substantial to almost perfect [Table S4]. Where the evaluated sections were distinguished according to the initial assessment, inter-observer concordance was moderate to substantial for all six items in gpHR; in gpHI this was almost perfect for five out of six items, with 'crypt architectural irregularities' as the only outlier with a moderate result. When the evaluated items were divided according to Mayo score, inter-observer concordance for all items of the RS was fair to substantial in the case of Mayo score 0 [range 0.35–0.69] and almost perfect in five out of six items [range 0.90–0.95] in the case of Mayo score 1, with again 'crypt architectural irregularities' as an outlier with a substantial result [Table S4].

3.8. Global visual score [GVS]

Overall inter-observer concordance in the GVS was substantial [0.74]. In those sections primarily assessed as gpHR this was moderate [0.52], whereas in biopsies scored as gpHI this was substantial [0.78]. When the evaluated sections were divided according to Mayo score, the inter-observer concordance for GVS was moderate in the case of Mayo score 0 [0.52] and almost perfect in the case of Mayo score 1 [0.92] [Table 3].

3.9. Intra-observer reproducibility histological activity scores

Intra-observer reproducibility of reader 1 [IN] and reader 2 [AT] in all scores is given in Table 4. Eighteen randomly selected sections [18/77; 23.4%] with 40 biopsies were blindly reassessed a second time by these two expert GI-pathologists. Overall, reader 1 showed the highest intra-observer reproducibility for the HGI with a kappa of 1 and with 100% agreement. For reader 2, the highest intra-observer reproducibility [0.86] was for GS without subscoring. With regard to the RS, reader 1 showed 100% agreement for

Downloaded from https://academic.oup.com/ecco-jcc/advance-article-abstract/doi/10.1093/ecco-jcc/jjx165/4733426
by University of Florida user
on 05 January 2018

Table 2. Histological reassessment of the Mayo 0 and 1 biopsies by expert GI-pathologists – % HR according to general pathologist [gpHR] vs expert GI-pathologists.

Mayo score	General pathologist	Expert GI-pathologists			
	gpHR*	HGI < 1†	GS < 0.1†	GS < 3.1†	RS 0†
Mayo 0 overall	89.6% [60/67]	90.0% [181/201]	29.4% [59/201]	91.0% [183/201]	25.4% [51/201]
Mayo 0 rectosigmoid	84.8% [28/33]	88.9% [88/99]	25.3% [25/99]	88.9% [88/99]	21.2% [21/99]
Mayo 1 rectosigmoid	20% [2/10]	70.0% [21/30]	16.7% [5/30]	63.3% [19/30]	20.0% [6/30]
Mayo 0 proximal colon	94.1% [32/34]	91.2% [93/102]	33.3% [34/102]	93.1% [95/102]	29.4% [30/102]

HGI = Harpaz Gupta Index; GS = Geboes score; RS = Riley score; *gpHR = biopsies initially evaluated as 'histological remission' by general pathologist; †common cut-offs for histological remission.

Table 3. Inter-observer concordance among GI-pathologists for each histological scoring index.

Index		Cronbach's α [95% CI]	Agreement [%]
Global Visual Scale [GVS]	Overall	0.74 [0.46–0.86]	8.0%
	gpHR*	0.52 [0.34–0.62]	9.8%
	gpHI**	0.78 [0.12–0.93]	0.0%
	Mayo 0†	0.51 [0.37–0.64]	9.2%
	Mayo 1††	0.92 [0.39–0.96]	0.0%
Harpaz Gupta Index [HGI]	Overall	0.61 [0.43–0.78]	76.6%
	gpHR*	0.51 [0.21–0.66]	82.5%
	gpHI**	0.63 [0.31–0.88]	50.0%
	Mayo †	0.55 [0.33–0.70]	80.6%
	Mayo 1††	0.64 [0.21–0.914	50.0%
Geboes Score [GS] with subscoring	Overall	0.84 [0.68–0.91]	11.7%
	gpHR*	0.65 [0.35–0.79]	14.3%
	gpHI**	0.91 [0.75–0.97]	0.0%
	Mayo 0†	0.70 [0.50–0.81]	13.4%
	Mayo 1††	0.93 [0.76–0.98]	0.0%
Geboes Score [GS] without subscoring	Overall	0.83 [0.67–0.90]	35.1%
	gpHR*	0.63 [0.34–0.78]	36.5%
	gpHI**	0.91 [0.76–0.97]	28.6%
	Mayo 0†	0.68 [0.46–0.80]	34.3%
	Mayo 1††	0.92 [0.74–0.99]	40.0%
Riley Score [RS]	Overall	0.91 [0.75–0.96]	14.3%
	gpHR*	0.73 [0.46–0.86]	15.9%
	gpHI**	0.95 [0.75–0.98]	7.1%
	Mayo 0†	0.79 [0.58–0.88]	14.9%
	Mayo 1††	0.96 [0.74–0.98]	10.0%

*Biopsies initially evaluated as 'histological remission' by general pathologist; **biopsies initially evaluated as 'histological inflammation' by general pathologist; †biopsies taken from colon segments assessed as endoscopic Mayo score 0, ††biopsies taken from colon segments assessed as endoscopic score Mayo 1]

'acute inflammatory cell infiltrate, crypt abscesses, mucin depletion and surface epithelial integrity'. Reader 2 showed similar results for 'crypt abscesses and surface epithelial integrity'.

3.10. Correlations between primary assessment [gp], endoscopic Mayo score and histological activity scores

Correlation between GS and GVS was substantial[R = 0.73], as was that between the RS and the GVS [R = 0.75]. Correlation between HGI and GVS was moderate [R = 0.41]. Correlation between the RS and the GS was almost perfect [R = 0.86]. Between the RS and the HGI, correlation was moderate [R = 0.44], as was that between HGI and GS [R = 0.56]. Correlation between the primary assessment and the Mayo score was moderate [R = 0.52]. There was no correlation between the primary assessment and the GVS [R = –0.04], the HGI [R = 0.03], the GS [R = 0.00] and the RS [R = –0.01] as scored by the expert GI-pathologists.

There was no correlation between the Mayo score and the GVS [R = –0.07], HGI [R = –0.03], GS [R = –0.11] and RS [R = –0.09] as scored by the expert GI-pathologists.

4.1. Discussion

This study documents poor concordance between 'daily practice' histological assessment of HR in UC biopsies by a general pathologist and expert reviews by GI-pathologists. This finding is important as our study mirrors daily practice where clinicians are struggling with the question whether to alter UC medication and the presence of HR may tip the balance in favour of a specific treatment strategy. We found strong correlations among expert GI-pathologists between the most frequently used histological activity scores for UC [RS and GS] and we observed a substantial inter-observer concordance for three histological scores [GS, HGI, RS], when re-evaluating UC biopsies assessed as HR by a general pathologist. Our study confirmed the reported excellent diagnostic properties of the most widely used UC histological disease activity indexes[24,25] specifically for UC colonic biopsies with MH and limited histological inflammation.

The significant discrepancy we found between histological assessment of UC biopsies in daily practice and expert GI-pathologists is probably explained by bias and the need for expertise. The GI-pathologists from our research panel were blinded for both clinical and endoscopic data, whereas the initial assessor [in clinical practice] had full access to clinical and endoscopy results. It is conceivable that the clinical information available to the general pathologist influenced the conclusions of the histological evaluation. Secondly, routine histological assessment of IBD mucosal biopsies obtained from a colon with MH is complex and probably requires expertise that goes beyond routine clinical review. Indeed, some IBD studies have shown considerable inter-observer variability between pathologists.[30,31] Histological inter-observer disagreement is by no means rare, nor is it limited to UC. In cancer treatment where results from pathology examinations are crucial, central pathology reading of biopsies taken in the context of trials is ubiquitous[32] and has made its way into daily oncology care.[33–35] This concept appears to have traction in IBD trials,[36] but is not yet in daily practice. We applaud the development of novel UC histological activity indexes[37–39] because we need accurate assessment of histological activity in both IBD trials and clinical practice. There are two other

Downloaded from https://academic.oup.com/ecco-jcc/advance-article-abstract/doi/10.1093/ecco-jcc/jjx165/4733426
by University of Florida user
on 05 January 2018

Table 4. Intra-observer reproducibility for each histological scoring index and separate items of the Riley score.

Index	Reader 1 [IN]	Reader 2 [AT]	Reader 1 [IN]	Reader 2 [AT]
	Cronbach's α [95% CI]	Cronbach's α [95% CI]	Agreement [%]	Agreement [%]
GVS	0.34 [-0.78–0.75]	0.85 [0.61–0.95]	55.6%	50%
HGI	0.00 [0.00–1.00]	0.61 [-0.05–0.90]	94.4%	88.9%
GS with subscoring	0.73 [0.27–0.90]	0.86 [0.63–0.95]	72.2%	55.6%
GS without subscoring	0.61 [-0.05–0.85]	0.86 [0.64–0.95]	88.9%	66.7%
RS	0.79 [0.44–0.92]	0.82[0.51–0.93]	61.1%	44.4%
RS item				
Acute inflammatory cell infiltrate	0.00 [0.00–1.00]	0.93 [0.82–0.98]	88.9%	94.4%
Crypt abcesses	0.00 [0.00–1.00]	1.00 [0.00–1.00]	94.4%	100%
Mucin depletion	0.00 [0.00–1.00]	-0.18 [-0.44–1]	77.8%	83.3%
Surface epithelial integrity	1.00 [0.00–1.00]	0.00 [0.00–1.00]	100%	94.4%
Chronic inflammatory cell infiltrate	0.79 [0.44–0.92]	0.78 [0.41–0.92]	94.4%	72.2%
Crypt architectural irregularities	0.93 [0.80–0.97]	0.79 [0.44–0.92]	72.2%	72.2%

GVS = Global Visual Scale; HGI = Harpaz Gupta Index; GS = Geboes score; RS = Riley score.

histological indexes for UC that merit discussion. Both the Nancy Index and the Robarts Histopathology Index [RHI] have shown good reproducibility, reliability and responsiveness, and underwent the most extensive validation of all existing histological activity indexes for UC.[37,38,40] The validation process for both indexes has been performed by expert pathologists, and both mainly use key items that are also present in existing scores such as GS, RS, HGI, Gramlich and GVS. These key items had high intra- and inter-observer agreement and were subsequently included in the Nancy Index and RHI. Given the overlap of the included histological items [such as ulceration and acute inflammatory cell infiltrate], inclusion of either the Nancy Index or the RHI in our study would not change our key message. The real matter at stake, based on current study results, is the gap between the initial and expert histological assessment. Efforts to overcome this gap should be directed at dedicated learning pathways and/or central reading facilities. The Nancy Index might gain a future central role in these learning pathways because it is simple and easy to use. We found higher inter-observer concordance between expert GI-pathologists in the biopsies that were assessed by a general pathologist as histological inflamed compared to HR for all three histological activity scores. Similar results were found after categorization into endoscopic Mayo score 0 and 1, in favour of the Mayo 1 group. Both findings suggest that is it more complicated, even for specialized GI-pathologists, to reach an agreement in cases of minor histological inflammation. This assumption concurs with a study from the 1990s that investigated the reliability of the interpretation of IBD colonic biopsies by specialized GI-pathologists. True normal biopsies were frequently assessed as 'possible' or 'likely' 'non-specific inflammation' without agreement on this item.[30]

In our study we found no correlation between the endoscopic Mayo score and any histological scores. In the case of Mayo 0 at endoscopy, we found histological activity in 20% [GS > 3.1] to 80% of cases [GS > 0.1] depending on the definition of HR used. In line with these results, a previous study described histological inflammation [GS ≥ 3.1] in 40% of cases with MH at endoscopy.[41] Unfortunately there is no unified definition of HR and the cut-off level of histological activity that is clinically relevant is still of debate.[15] Our study has several strengths and clinical implications. To mimic daily practice with regard to HR decision-making, we only used biopsies reported as endoscopic MH, and we involved expert GI-pathologists from both academic and non-academic hospitals. For these histology analyses, we used blinded readers, which reduces the potential bias of taking the clinical condition of the patient into consideration. Most importantly, the described discrepancy between histological assessment of UC biopsies in daily practice and reassessment by dedicated GI-pathologists may have important implications in daily practice. Medical decision-making may increasingly depend on histological remission, and therefore it is important to realize and act upon this finding. There are some limitations of this study. First, the retrospective design of the study harbours the risk of bias. Secondly, the consensus meeting between GI-pathologists, before initiation of the study, may have caused bias by achieving a better inter-observer agreement. Third, we did not use central reading[42] for the endoscopic Mayo score, but used the Mayo score as reported by the endoscopist. In addition, the Mayo score has a variable inter-observer concordance,[42–44] although this is likely to have little impact on the scope of this article.

In summary, the use of existing histological activity indexes evaluating UC colonic biopsies that were initially assessed as HR resulted in substantial inter-observer concordance. Of alarm is the observation that there was no correlation between the primary assessment of UC biopsies by the general pathologist and the reassessment by blinded expert GI-pathologists. This may have important implications for the selection process of a unified histological disease activity score in UC, and for the implementation of HR as a UC treatment target in daily practice.

Funding

None.

Conflict of Interest.

The authors have no conflicts of interest to declare.

Acknowledgements

We would like to thank Dr Wiebe Pestman for his statistical advice during the writing of this article.

Author Contributions

TR: study concept and design; acquisition of data; analysis and interpretation of data; statistical analysis, drafting of the manuscript. PM: acquisition of

Downloaded from https://academic.oup.com/ecco-jcc/advance-article-abstract/doi/10.1093/ecco-jcc/jjx165/4733426
by University of Florida user
on 05 January 2018

data; analysis and interpretation of data; statistical analysis, drafting of the manuscript. AvT: technical support, analysis of data, critical revision of the manuscript for important intellectual content. CB: technical support, analysis of data, critical revision of the manuscript for important intellectual content. IN: technical support, analysis of data, critical revision of the manuscript for important intellectual content. JPHD: critical revision of the manuscript for important intellectual content; study supervison. FH: critical revision of the manuscript for important intellectual content supervision, study supervision.

Supplementary Data

Supplementary data are available at *ECCO-JCC* online.

References

1. 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.
2. van Dullemen HM, van Deventer SJ, Hommes DW, *et al*. Treatment of Crohn's disease with anti-tumor necrosis factor chimeric monoclonal antibody (cA2). *Gastroenterology* 1995;**109**:129–35.
3. Rutgeerts P, D'Haens G, Targan S, *et al*. Efficacy and safety of retreatment with anti-tumor necrosis factor antibody (infliximab) to maintain remission in Crohn's disease. *Gastroenterology* 1999;**117**:761–9.
4. Hanauer SB, Feagan BG, Lichtenstein GR, *et al*.; ACCENT I Study Group. Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial. *Lancet* 2002;**359**:1541–9.
5. Sands BE, Anderson FH, Bernstein CN, *et al*. Infliximab maintenance therapy for fistulizing Crohn's disease. *N Engl J Med* 2004;**350**:876–85.
6. Rutgeerts P, Diamond RH, Bala M, *et al*. Scheduled maintenance treatment with infliximab is superior to episodic treatment for the healing of mucosal ulceration associated with Crohn's disease. *Gastrointest Endosc* 2006;**63**:433–42; quiz 464.
7. Walsh A, Palmer R, Travis S. Mucosal healing as a target of therapy for colonic inflammatory bowel disease and methods to score disease activity. *Gastrointest Endosc Clin N Am* 2014;**24**:367–78.
8. Theede K, Kiszka-Kanowitz M, Nordgaard-Lassen I, Mertz Nielsen A. The impact of endoscopic inflammation and mucosal healing on health-related quality of life in ulcerative colitis patients. *J Crohns Colitis* 2015;**9**:625–32.
9. Barreiro-de Acosta M, Lorenzo A, Mera J, Dominguez-Munoz JE. Mucosal healing and steroid-sparing associated with infliximab for steroid-dependent ulcerative colitis. *J Crohns Colitis* 2009;**3**:271–6.
10. Barreiro-de Acosta M, Vallejo N, de la Iglesia D, *et al*. Evaluation of the risk of relapse in ulcerative colitis according to the degree of mucosal healing (Mayo 0 vs 1): a longitudinal cohort study. *J Crohns Colitis* 2016;**10**:13–9.
11. 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.
12. D'Haens G, Baert F, van Assche G, *et al*.; Belgian Inflammatory Bowel Disease Research Group; North-Holland Gut Club. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. *Lancet* 2008;**371**:660–7.
13. 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.
14. 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.
15. Bryant RV, Winer S, Travis SP, Riddell RH. Systematic review: histological remission in inflammatory bowel disease. Is 'complete' remission the new treatment paradigm? An IOIBD initiative. *J Crohns Colitis* 2014;**8**:1582–97.
16. 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.
17. Travis SP, Higgins PD, Orchard T, *et al*. Review article: defining remission in ulcerative colitis. *Aliment Pharmacol Ther* 2011;**34**:113–24.
18. Korelitz BI. Mucosal healing as an index of colitis activity: back to histological healing for future indices. *Inflamm Bowel Dis* 2010;**16**:1628–30.
19. Villanacci V, Antonelli E, Geboes K, Casella G, Bassotti G. Histological healing in inflammatory bowel disease: a still unfulfilled promise. *World J Gastroenterol* 2013;**19**:968–78.
20. Peyrin-Biroulet L, Bressenot A, Kampman W. Histologic remission: the ultimate therapeutic goal in ulcerative colitis? *Clin Gastroenterol Hepatol* 2014;**12**:929–34.e2.
21. 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.
22. Riley SA, Mani V, Goodman MJ, Dutt S, Herd ME. Microscopic activity in ulcerative colitis: what does it mean? *Gut* 1991;**32**:174–8.
23. Gupta RB, Harpaz N, Itzkowitz S, *et al*. Histologic inflammation is a risk factor for progression to colorectal neoplasia in ulcerative colitis: a cohort study. *Gastroenterology* 2007;**133**:1099–105; quiz 1340–1.
24. Mosli MH, Feagan BG, Zou G, *et al*. Reproducibility of histological assessments of disease activity in UC. *Gut* 2015;**64**:1765–73.
25. 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.
26. Satsangi J, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut* 2006;**55**:749–53.
27. 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**:1625–9.
28. Yokoyama K, Kobayashi K, Mukae M, Sada M, Koizumi W. Clinical study of the relation between mucosal healing and long-term outcomes in ulcerative colitis. *Gastroenterol Res Pract* 2013;**2013**:192794.
29. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74.
30. Theodossi A, Spiegelhalter DJ, Jass J, *et al*. Observer variation and discriminatory value of biopsy features in inflammatory bowel disease. *Gut* 1994;**35**:961–8.
31. Farmer M, Petras RE, Hunt LE, Janosky JE, Galandiuk S. The importance of diagnostic accuracy in colonic inflammatory bowel disease. *Am J Gastroenterol* 2000;**95**:3184–8.
32. Gilchrist KW, Harrington DP, Wolf BC, Neiman RS. Statistical and empirical evaluation of histopathologic reviews for quality assurance in the Eastern Cooperative Oncology Group. *Cancer* 1988;**62**:861–8.
33. Ghadjar P, Hayoz S, Genitsch V, *et al*.; Swiss Group for Clinical Cancer Research (SAKK). Importance and outcome relevance of central pathology review in prostatectomy specimens: data from the SAKK 09/10 randomized trial on prostate cancer. *BJU Int* 2016;**120**:E45–E51.
34. Kishimoto R, Saika T, Bekku K, *et al*. The impact of pathologic review by central pathologist on selection for treatment modality of localized prostate cancer in candidate for brachytherapy. *World J Urol* 2012;**30**:375–8.
35. Soga N, Yatabe Y, Kageyama T, Ogura Y, Hayashi N. Review of bioptic Gleason scores by central pathologist modifies the risk classification in prostate cancer. *Urol Int* 2015;**95**:452–6.
36. Magro F, Lopes SI, Lopes J, *et al*.; Portuguese IBD group [GEDII]. Histological outcomes and predictive value of faecal markers in moderately to severely active ulcerative colitis patients receiving infliximab. *J Crohns Colitis* 2016;**10**:1407–16.
37. 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.
38. Mosli MH, Feagan BG, Zou G, *et al*. Development and validation of a histological index for UC. *Gut* 2017;**66**:50–8.
39. Jauregui-Amezaga A, Geerits A, Das Y, *et al*. A simplified Geboes score for ulcerative colitis. *J Crohns Colitis* 2017;**11**:305–13.
40. 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.

Downloaded from https://academic.oup.com/ecco-jcc/advance-article-abstract/doi/10.1093/ecco-jcc/jjx165/4733426
by University of Florida user
on 05 January 2018

41. 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.
42. Feagan BG, Sandborn WJ, D'Haens G, *et al*. The role of centralized reading of endoscopy in a randomized controlled trial of mesalamine for ulcerative colitis. *Gastroenterology* 2013;**145**:149–157.e2.
43. Daperno M, Comberlato M, Bossa F, *et al*. Inter-observer agreement in endoscopic scoring systems: preliminary report of an ongoing study from the Italian Group for Inflammatory Bowel Disease (IG-IBD). *Dig Liver Dis* 2014;**46**:969–73.
44. Osada T, Ohkusa T, Yokoyama T, *et al*. Comparison of several activity indices for the evaluation of endoscopic activity in UC: inter- and intraobserver consistency. *Inflamm Bowel Dis* 2010;**16**:192–7.

Downloaded from https://academic.oup.com/ecco-jcc/advance-article-abstract/doi/10.1093/ecco-jcc/jjx165/4733426
by University of Florida user
on 05 January 2018