# A multimodal (FACILE) classification for optical diagnosis of inflammatory bowel disease associated neoplasia

Iacucci, M.; McQuaid, K.; Gui, X.S.; Iwao, Y.; Lethebe, B.C.; Lowerison, M.; Matsumoto, T.; Shivaji, U.N.; Smith, S.C.L.; Subramanian, V.; Uraoka, T.; Sanduleanu, S.; Ghosh, S.; Kiesslich, R.

*Document Version*
Peer reviewed version

Link to publication on Research at Birmingham portal

"A multimodal (FACILE) classification for optical diagnosis of inflammatory bowel disease associated neoplasia".

Iacucci M[1,2], McQuaid K[3], Gui XS [1,] Iwao Y[9] , Lethebe  BC[7] Lowerison M[7], Matsumoto T[5] , Shivaji U [2] , Smith S [2 ,] , Subramanian V[6] , Uraoka T[4], Sanduleanu S [8] ,  Ghosh S[1,2] ,  Kiesslich R[9]


[1]Division of Gastroenterology, University of Calgary, Canada,
[2] Institute Translational of Medicine, Institute of immunology and immunotherapy and NIHR Biomedical Research Centre, University of Birmingham, UK
[3]Division of Gastroenterology, University of California San Francisco, USA
[4] Division of Gastroenterology, Tokyo Medical Center, Japan
[5] Division of Gastroenterology, Iwate Medical University, Japan
[6] Division of Gastroenterology, University of Leeds, UK
[7]Clinical Research Unit, University of Calgary, Canada
[8]Division of Gastroenterology and Hepatology, Maastricht University Medical Center, The Netherlands
[9]Center for Preventive Medicine, Keio University I, Tokyo, Japan
[9]Division of Gastroenterology, HSK Hospital, Wiesbaden ,Germany


Correspondence:
Marietta Iacucci MD, PhD, FASGE
Reader/Senior Associate Professor of Gastroenterology
Institute of Translational Medicine
University of Birmingham, United Kingdom

Adjunct Clinical Associate Professor of Medicine
University of Calgary, Canada

Institute of Translational Medicine
Heritage Building for Research & Development
University Hospitals Birmingham NHS Foundation Trust
Edgbaston B15 2TT, Birmingham, UK
Telephone +44 (0)1213718119
email: m.iacucci@bham.ac.uk

GS and KR contributed equally as last/senior author of this manuscript

**Conflicts of Interest**

MI: received unrestricted research grant from Pentax USA

SS: unrestricted research grant Pentax Europe

None of the other authors have any conflict of interest to declare

**Abstract**

Introduction

Characterisation of colonic lesions in inflammatory bowel disease (IBD) remains challenging. We developed an endoscopic classification of visual **characteristics of** imaging **for colonic lesions** to **identify colitis associated neoplasia** using multimodal advanced endoscopic imaging {Frankfurt Advanced Chromoendoscopic Ibd LEsions (FACILE)}.

Methods

The study was conducted in 3 phases: 1) Development. An expert panel defined endoscopic signs and predictors of dysplasia in IBD using multivariable logistic regression creating the FACILE classification. 2) Validation. Using 60 IBD lesions from library (dysplasia, cancer, SSA/Ps, inflammatory polyps), we performed 2 assessments of diagnostic accuracy for neoplasia and inter-observer agreement between experts using FACILE. 3) Reproducibility. We tested reproducibility of the FACILE in gastroenterologists, trainees and Jiunior doctors after a training module.

Results

The experts initially selected criteria such as morphology, colour, surface, vessels architecture, sign of inflammation, border of lesion.
Multivariable logistic regression confirmed that non-polypoid lesion, irregular vessel architecture, irregular surface pattern and signs of inflammation within the lesion were predictors of dysplasia. Area under the curve of this logistic model using a bootstrapped estimate was 0.76 (0.73-0.78).

The training module resulted in improvement in accuracy and kappa agreement in all non-experts, though in trainees and junior doctors the kappa agreement was still moderate and poor respectively.

Conclusion

We developed, validated and reproduced a new endoscopic classification (FACILE) using all imaging modalities for diagnosis of dysplasia in IBD. Flat shape, irregular surface, vascular pattern and sign of inflammation predicted dysplasia. The diagnostic performance of all non-expert participants improved after a training module.

Key words: Surveillance colonoscopy, Neoplastic lesions in IBD, Virtual Electronic Chromoendoscopy, FACILE

**Abbreviation**

IBD =Inflammatory Bowel Disease

SCENIC = Surveillance for Colorectal Endoscopic Neoplasia Detection and Management in Inflammatory Bowel Disease Patients: International Consensus Recommendations

SSA/Ps=Sessile Serrated adenoma/Polyps

HD=High Definition

VCE=Virtual Electronic Chromoendoscopy

DCE=Dye Chromoendoscopy

NBI=Narrow Banding Imaging

LGD=Low grade dysplasia

HGD=High grade dysplasia

CL=Colonic Lesions

**Introduction**

SCENIC (Surveillance for Colorectal Endoscopic Neoplasia Detection and Management in Inflammatory Bowel Disease Patients: International Consensus Statement) identified the endoscopic techniques that need to be adopted to increase the detection rate of dysplasia in IBD. SCENIC also introduced a new terminology  to describe the shape of colonic lesions, a new concept of endoscopic resectablility and the terminology underpinning the morphology of the colonic dysplastic lesions but did not consider the endoscopic  features  to predict histology and invasiveness of the lesions [1]. The endoscopic classification of morphology and characterization for colonic lesions used are Paris classification, Kudo pit pattern, and Hazewinkel criteria for sessile serrated adenoma/polyp (SSA/Ps). [2-5]

However, it is controversial whether the Kudo pit pattern can also be used to predict histology of colonic lesions in IBD especially when assessed by standard colonoscopes without magnification [6-9]. Regenerative changes in IBD may masquerade as dysplastic lesions showing Kudo pit pattern IIIL and IV) and hence Kudo pit pattern may be misleading. It has been demonstrated that with the new generation of High Definition (HD) with or without virtual chromoendoscopy (VCE) and without magnification, lesions associated with IBD can be assessed by endoscopic features.  New advanced technologies have made unmasking of dysplasia easier. [10-13] Recently, Sugimoto *et al* have classified the morphologic features of High Grade Dysplasia (HGD) using the SCENIC guidelines. The findings of flat/superficial-elevated area, red discoloration and left colon localization were associated with HGD in IBD patients. The magnifying colonoscopes were also useful in distinguishing lesions from the surrounding mucosa and helped  therapeutic endoscopic management. [14]

In our study, we aimed to develop an endoscopic classification of visual characteristics of advanced imaging for colonic lesions to identify and validate

colitis associated neoplasia by international expert consensus conference at Frankfurt (Frankfurt Advanced Chromoendoscopic IBD LEsions classification - FACILE).  As all lesions were in colitic areas and were dysplasia, cancer or SSA/P .The endoscopic visual characteristics aimed to distinguish neoplastic lesions from non-neoplastic lesions.  We further evaluated the reproducibility and reliability of the FACILE classification in the setting of experienced gastroenterologists/trainee gastroenterologists and junior doctors not exposed to endoscopy using a library of images of colonic lesions in IBD.

**Methods**

The Calgary Conjoint Health Services Research Ethics Board of the University of Calgary (CHREB) approved the study (REB13-0960).

We conducted a multistep study to develop and validate classification of IBD lesions found at surveillance using different advanced endoscopic techniques such as HD White Light Endoscopy (WLE) Dye Chromoendoscopy (DCE), VCE including iSCAN, NBI.. The study used image library and was structured and conducted in 3 phases (Figure 1):

*Phase 1*) Development. A panel of international experts, through roundtable iterative discussion, in a modified Delphi method to achieve consensus, defined endoscopic signs of colitis associated neoplasia including dysplastic and  SSA/P and cancer {versus non-neoplastic lesions} using international nomenclature (SCENIC guidelines, Paris endoscopic classification, Kudo pit-pattern classification and Hazewinkel criteria [1-5] (Supplementary Figure 2). Predictors of neoplasia were determined using univariable and multivariable logistic regression models to determine the strength of endoscopic predictors of neoplastic Vs. non-neoplastic lesions diagnosis by histology. (Table 1.) This led to refinement and final FACILE classification system shown in Figure 3.

6

*Phase 2*) <u>Validation</u>. We performed 2 consecutive measurements on different days of the diagnostic accuracy for dysplasia and inter-observer agreement between experts using the FACILE classification. A second presentation with the same pictures presented in a different random order and on the second day of the consensus to minimize recall bias, was provided to all participants. (Table 2.) Inter-observer agreement was calculated using Fleiss' Kappa, and the confidence intervals were established using a bootstrap approach.

*Phase 3*) <u>Reproducibility</u>. We tested the reproducibility of the FACILE classification in experienced (consultant) gastroenterologists, trainees and junior doctors not exposed to endoscopy using a computerized training module. (Containing a different set of images and video clips) before and after training. They then assessed the same 60 pictures of colonic lesions that had been assessed previously by the international experts.  (Table 3-4, Figure 4)

At all stages, the raters were blinded to the nature of images used in the study.

**Image-Library**

 <u>Source of data</u>: A library of 60 high quality still images of colonic lesions (20 HD, 20 DCE and 20 VCE-each distinct lesion) was used by the international experts for creating the FACILE classification. Of these, 33 (55%) were dysplasia (12 HD, 10 DCE ,11 VCE);  6 (10%) cancer (2 HD, 2 DCE, 2 VCE);  9 (15%) SSA/Ps (3 HD, 2 DCE, 4 VCE); and 12 (20%) inflammatory polyps (3 HD, 6 DCE, 3 VCE). The 12 images representing benign lesions were variable and had a spectrum of ulcerated, non-ulcerated, polypoid and non-polypoid lesions. The anonymized images selected for the study were collected by MI, RK and TU referred for surveillance colonoscopy using HD, DCE and VCE with iSCAN (7000EPKi, Pentax, Japan), and NBI (EVIS Lucera System CLV 260 from Japan and UK; Olympus Tokyo, Japan). Two experienced gastroenterologists

in advanced technologies (KR and MI) selected from a library of more than 200 pictures, 60 high quality pictures that reflected all the different histological categories.  The participants were not aware of the distribution of lesions in the library.

**Participants**

Four groups were involved in this study in different phases (a) expert international gastroenterologists experienced in advanced endoscopic technologies for a median of 15 years  (n= 3: Europe, n=3: USA and n=2: Japan) and had performed more than 8000 colonoscopies each and all of them were familiar with DCE, NBI, iSCAN and FICE (b) experienced consultant gastroenterologists (n=5: Europe). They were in practice for a median of 14 years and had performed an average of 3900 colonoscopies (range 3000 - 4500); c) gastroenterology trainees (n=8: Europe) were in training for a median of 4 years and had performed an average of 300 colonoscopies (range 160-600). All were exposed to DCE and VCE with NBI and iSCAN d) junior doctors naïve to endoscopy (n=6: Europe) who have never been exposed to endoscopy.

**Outcomes:**

To develop and validate a classification (FACILE) that can distinguish colitis associated neoplastic from non-neoplastic lesions without using Kudo pit patterns in IBD surveillance colonoscopy and reproduce the classification in a non-expert cohort.

**First Phase of the study:**
**Development of FACILE classification by international expert gastroenterologists**

Four expert gastroenterologists (MI, SS, TU, MT) presented Microsoft PowerPoint slides of SCENIC guidelines and of colonic IBD lesions( photos and videos) detected by DCE, VCE NBI and VCE iSCAN.

The international experts through discussion and in a stepwise feedback fashion (to ensure equal participation) defined the endoscopic findings of the FACILE classification [1-5]. They selected endoscopic criteria for the classification that could be visualized by all the available advanced imaging techniques such as morphology, surface, vessels architecture and border of the lesions. The experts then decided not to integrate Kudo pit pattern terminology to avoid uncertainty in the absence of magnification and in presence of changes due to healing and challenges of interpreting regenerative changes.

We performed univariable and multivariable logistic regression using the endoscopic findings defined by experts to predict histology (neoplastic vs non-neoplastic) and generate the final FACILE classification (Supplementary Figure 2, Figure 3 and Table 1)

**Second phase of the study – Validation of FACILE classification**
All eight participants performed 2 consecutive assessments, on different days, of the diagnostic accuracy for dysplasia and inter-observer agreement using the FACILE classification. A total of 60 high quality pictures (20 HD, 20 DCE and 20 VCE) were projected to the participants. These slides were different from the materials used for development of FACILE. The participants were not aware of the distribution of lesions in the library.

A second presentation with the same pictures was presented in a different random order on the second day of the consensus in order to minimize recall bias. Each endoscopist individually scored each of the criteria of the FACILE

classification (morphology, colour, surface pattern, vessel pattern, demarcation and inflammation within or surrounding the lesion), predicted histology as neoplastic (dysplasia, and SSA/Ps), cancer and inflammatory polyps and assigned a level of confidence to the prediction of histology (high Vs. low). (Figure 1).

.

**Third phase of the study- Reproducibility of the FACILE classification after computerized training module**

An interactive Microsoft PowerPoint teaching module was presented by one endoscopist (MI). The presentation included both slides and videos of colonic lesions collected during surveillance colonoscopy in IBD, including 50 pictures and 10 videos (each video of the length of 60 seconds) of colonic lesions including SSA/Ps, polypoid and non-polypoid dysplasia,  cancer and inflammatory polyps using different modalities (HD-WLE, DCE, NBI, iSCAN) . The contents of the module also included an introduction to the SCENIC guidelines, Paris classification, Kudo pit pattern, Hazewinkel criteria for SSA/Ps and the FACILE classification [1-5]. The contents were distinct from phase 1 and phase 2 of the study and this phase reflected an educational training Microsoft Powerpoint teaching module.

The participants' ability to categorize the colonic lesions and predict neoplastic lesions before and after the training (validation) was determined.

**Histopathological assessment**

Histopathology of the colonic lesions were used as the gold standard. The pathology reports were analysed by two pathologists as part of quality assurance of accreditated laboratory services for standard clinical care. Neoplastic changes were classified as SSA, traditional serrated adenomas,

sporadic tubular adenoma, inflammatory lesions, LGD, HGD or adenocarcinoma in the colitic areas, according to the modified Vienna classification. [15]

**Statistical analysis and sample size**

The results were transferred to a Microsoft Excel database and exported to STATA Version 13.1; Statistical analysis was conducted using the R Statistical Software (R version 3.4.4).

Exploratory univariable logistic regression analysis was performed for selecting endoscopic variables associated with the presence of dysplasia, SSA/P or cancer. Those features that were significant at the univariable stage (At the univariable stage, we considered variables to be significant if the p-value was less than 0.05) were included in a multivariable logistic regression, with no interaction terms.

60 pictures provided 80% power to detect a kappa agreement difference from 0.40 to 0.60 (moderate –good agreement using a two—sided significance level of 0.05. (16). The inter-observer agreement was calculated by using the Fleiss kappa coefficients (>6 observers; >3 categories) [ 16-17] and confidence intervals were established using a bootstrap method. Sixty images provided substantial inter-observer agreement (Fleiss kappa coefficient: 0.80, 95% CI, 0.70-0.90) in order to predict lesion histology in the surveillance of IBD patients [17]. We performed sensitivity analyses by systematically excluding reviewers and recalculating the Kappa statistic to obtain a robust estimate of agreement.

A sample size of non-experts and trainees of 1140 observations (60 pictures 19 reviewers, 5 consultants, 8 registrars and 6 trainees) was calculated to be

statistically adequate to externally reproduce and validate the FACILE
IBD classification.

Sensitivity, specificity, PPV, NPV, and accuracy with their 95% confidence
intervals were calculated before and after training. p-values comparing pre-
and post-test accuracies were calculated using a mixed-modelling or random
intercept logistic regression to account for the repeated reviewers. A binary
indicator was created for each observation in the pre-test and post-test that
indicated correct and incorrect assessments. Then a random-intercept model
was fit using the pre-post indicator as the predictor variable, and an
adjustment for reviewer. The p-value was derived from the resulting z-statistic
for the coefficient from the model.

## Results

**Model Development of new classification (phase 1) -**
Figure 1 shows the first FACILE classification developed by experts and the
comparison of diagnostic performance for colonic lesions in IBD across the
international experts is shown in Table 2.

**Model performance by Multivariable analysis of endoscopic features
to predict histology of colonic lesions (neoplastic vs non-neoplastic)
(phase 1)**
Multivariable analysis confirmed that the endoscopic findings of the
morphology of non-polypoid lesion OR 3.13 (95% CI: 1.32 -7.25), irregular
vessel architecture OR 3.49 (95% CI: 1.74 -7.10), sign of inflammation within
the lesion OR 2.42 (95% CI: 1.24 – 4.79), irregular surface pattern OR 8.89
(95% CI: 3.21 – 25.96) were predictors of neoplasia.  Multivariable analysis of
performance characteristics of individual and combined endoscopic criteria to

predict dysplastic lesions associated with colitis when used by expert endoscopists is detailed in table 1.

Other parameters such as resectability and demarcation of lesions did not contribute to prediction of neoplasia.

The sensitivity, specificity, at the multivariable analysis stage were 94% (95% CI: 90-96%), 51%( 95% CI: 43-58%) for prediction of neoplasia and the area under the curve (bootstrap AUC) of this logistic model using a bootstrapped estimate was 0.76 (0.73-0.78).

**Validation and diagnostic accuracy for international experts. (phase 2)**

Among the international experts sensitivity, specificity, and accuracy in predicting neoplastic histology, were similar on the first day (72%, 74%, 72%) and on the second day (72%, 73%, 72 %). Individual rater accuracy ranged from 67% (95%CI 53 to 78) to 78% (95%CI 65 to 87). Sensitivity, specificity, accuracy, for predictions made with high confidence were 72%, 90%, 76%, which were significantly more accurate compared with a low confidence of diagnosis in the post test (76% vs 65%; P= 0.01). (Table 3).

**Reproducibility and inter-observer agreement among experienced, trainee gastroenterologists and naïve junior doctors (phase 3).**

The experienced gastroenterologists showed an improvement in the sensitivity, accuracy, and proportion of high- confidence diagnoses for characterization of IBD colonic lesions in the post training compared with the pre training. Their performance characteristics in the pre Vs post training were: sensitivity 81% Vs 90% (P =0.009), specificity 68% Vs 56% (P =0.34), accuracy 79% Vs 86% (P =0.01), and accuracy of prediction in the high-confidence diagnoses was 82% Vs 85%.

The trainee gastroenterologists also showed an improvement in their diagnostic performance in the pre Vs post training which was: sensitivity 80% Vs 82% (p=0.78) specificity 54% Vs 78% (p=0.02) and accuracy 75% Vs 81% (p=0.09),

Finally, the junior doctors (naïve to endoscopy) showed an improvement in pre Vs post training module. The sensitivity was 48% Vs 80% (p<0.001) specificity 44% Vs 58%( p=0.19) and accuracy 48% Vs 77% (p<0.001) (Table 2, Figure. 4)

Inter-observer agreement of the experienced gastroenterologists was good - this improved from the pre-training (kappa= 0.41, 95% CI: 0.23- 0.59) to post training (kappa = 0.67, 95% CI :0.46-0.82).

Inter-observer agreement of the trainees gastroenterology was moderate and improved from the pre-training (kappa =0.30;  95% CI: 0.22-0.44)) to post training (kappa = 0.41, 95% CI:0.28-0.58).

Amongst the junior endoscopy naïve doctors despite their improvement in the pre-training  (kappa=0.15, 95% CI 0.06-0.27) to post training  (kappa=0.20, 95% CI 0.09-0.33) the inter-observer agreement was poor  (Table 4).

Among this group of non-expert gastroenterologists, the sensitivity, specificity, PPV, NPV and accuracy for diagnosis of dysplasia associated with colitis using all four predictors were 30%, 96%, 96%, 29%, and 45% respectively (compared with 65%, 81%, 86%, 57% and 71% respectively in expert gastroenterologists shown in table 1).  Likewise, the accuracy for diagnosis of dysplasia using three, two or one predictors was also lower for this group of non-expert gastroenterologists (data not shown) (when compared with expert gastroenterologists table 1).

The proportion of high confidence diagnosis from FACILE IBD was dependent on experience; highest for international experts, and the lower for experienced gastroenterologists and still lower for trainees and lowest for endoscopy naïve junior doctors – training module administration did not consistently improve the proportion of high confidence diagnosis. (Table 3)

**Discussion**

We have developed a new endoscopic classification system to characterize neoplastic lesions in IBD. A special focus of the international experts was on endoscopic detailed features that could only be visualized by using advanced imaging techniques (DCE, NBI, iSCAN). We created the FACILE classification through a consensus and did a univariable (data not shown) followed by multivariable analysis to identify the endoscopic findings that were predictive of histology. We standardized the endoscopic features of colonic lesions in IBD using advanced endoscopic technologies without magnification to predict histology. [13,14,18)

Multivariable analysis showed that morphology, surface and vessel parameters are strong predictors and were integrated in the simplified FACILE classification. (Table 1) The irregular vessel architecture and non–polypoid lesion morphology were the best endoscopic predictors of dysplastic lesions determined by histology. Other endoscopic predictors of dysplastic lesions were irregular surface pattern and signs of inflammation within the lesions but less strong predictors.

Simplified FACILE classification allowed *in vivo* diagnosis of dysplasia with high sensitivity, specificity and diagnostic accuracy. This was true for experts – however also novice, trainees and experienced gastroenterologists could improve their diagnostic performance using FACILE.

This suggest that the FACILE classification is promising and may help in management decisions to perform local resections at colonoscopy or proceed to colectomy. The diagnostic accuracy of FACILE classification was 85%, nearly similar to >90% showed in the validation of the NICE classification for colonic polyps using the NBI technologies. [19]

The value of Kudo pit pattern to predict histology remains controversial in IBD patients especially when these lesions are assessed by using standard scopes without magnification. The colonic mucosa of IBD patients might be distorted due to long standing chronic inflammation and regenerative changes; furthermore dye spraying may also obscure Kudo pit pattern [9]. Recently, Carballal *et al* in routine clinical practice of DCE showed by univariable statistical analysis four findings: location at proximal colon, protruding morphology (Paris 0–Ip and 0–Is), loss of innominate lines and neoplastic pit pattern (IIIs, IIIL, IV and V) were predictive of dysplasia in IBD. [7] We also confirmed in our randomized surveillance study that the Kudo pit pattern (IIO, III-V) was an important predictive feature of dysplasia. [8] Recently, Bisschops *et al* assessed the accuracy levels of agreement amongst experts of Kudo pit pattern in UC with non-magnified NBI (6). However, because of the controversy, the international experts agreed not to consider the Kudo pit pattern in the FACILE classification when using technologies without magnification.

We also showed that a computerized structured interactive training module using FACILE amongst gastroenterologists and trainees improved performance.[21-23]
The accuracy of performance of junior 'endoscopy naïve' doctors improved from 48% in the pre-test to 77% in the post-test (P<0.001). (Figure 4)
Another impact of the training module was seen in the significant increment in

the proportion of high-confidence diagnoses of junior endoscopy naïve doctors 79% in the in the post-test compared with the pre-test of 37% (p<0.0001) (table 3).

These results added value of the FACILE classification system to existing characterization methods such as NICE and WASP in non-colitic patients as advanced technologies will be increasingly adopted in future.[5,19]

A strength of this study is the initial design of the FACILE classification by international experts by a consensus, selection of criteria by multivariable analysis relating best to prediction of dysplasia followed by involvement of experienced gastroenterologists, trainees and junior doctors pre- and post administration of a computerized training module. The process led to creation of a simplified FACILE classification that can be adopted by gastroenterologists after training by a computerized training module.

Limitations: Our study has certain limitations, as we have included high quality of still pictures rather than video clips that can reproduce a real-life assessment. In this study we do not try to compare different platforms or techniques (ie VCE-NBI, VCE-iSCAN, DCE) but develop a common platform that may then be applied to real life studies using different platforms.
We did not use confounders and mediators in order to develop our predictive modelling as these were not clear in our imaging characteristics based prediction of histology. We attempted alternative modelling and the resulting model from the LASSO regression using a cross fold validation technique was not meaningfully different than the models we have shown here.

The diagnostic accuracy for neoplasia associated with colitis remained superior for expert gastroenterologists (table 1) compared with the phase 3 group,

17

mentioned in 'Results'. The Kappa agreement in trainee gastroenterologists was moderate and junior doctors fair. A more extensive training module and pre-training will be tested in future to improve agreement in these more junior doctors. In future, a different group of experts who did not design FACILE will be important for further information about operating characteristics using a wider spectrum of lesions.

We did not include hyperplastic polyps in the image set. Many hyperplastic polyps occur in non-colitic areas and therefore are distinguished by their characteristics (such as Kudo pit pattern I or II) from neoplastic lesions . However the focus of FACILE is all lesions in colitic areas. We did not use the Kudo pit pattern in this study and consequently, we did not include hyperplastic polyps from colitic areas and left at this stage the diagnosis of such polyps to histology.  We acknowledge it is a limitation.

Lastly, overall accuracy is dependent on prevalence and hence the low prevalence of 'benign' images is a limitation. However, the prevalence of lesions at surveillance colonoscopy can also be very variable. Larger numbers of images will be used in the real- life study that is planned.  However the raters had no knowledge of the prevalence.

Implications: We hope the classification will help evaluation of lesions detected at surveillance in colitic areas as dysplastic (neoplastic) vs non-neoplastic and take real time decisions regarding whether it is necessary and feasible to resect lesions showing neoplastic features, without waiting for biopsies.

Our study will support further development of optical characterization of lesions in IBD, so that real life prospective assessment can be evaluated in studies.

Figure 1: Flow chart of study design

Supplementary Figure 2: First Classification and Endoscopic form used for scoring of Frankfurt Advanced Chromoendoscopic Ibd LEsions (FACILE) classification used by each participants

Figure 3:  Final Classification -Frankfurt Advanced Chromoendoscopic Ibd Lesions (FACILE) classification

Figure 4: Comparison of the diagnostic accuracy of each group before and after participation in a training computerized module about the use of the FACILE classification.

**References**

1.Laine L, Kaltenbach T, Barkun A, et al. SCENIC international consensus statement on surveillance and management of dysplasia in inflammatory bowel disease. Gastroenterology. 2015; 148: 639-651.

2. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002. Gastrointestinal Endoscopy. 2003; 58:S3-43.

3.Kudo S, Rubio CA, Teixeira CR, et al. Pit pattern in colorectal neoplasia: endoscopic magnifying view. Endoscopy. 2001; 33:367-373.

4.  Hazewinkel Y, López-Cerón M, East JE, et al Endoscopic features of sessile serrated adenomas: validation by international experts using high-resolution white-light endoscopy and narrow-band imaging. Gastrointest Endosc 2013; 77: 916-24

5. IJspeert JE, Bastiaansen BA, van Leerdam ME et al. Development and validation of the WASP classification system for optical diagnosis of adenoma, Hyperplastic polyps and sessile serrated adenoma / polyps. Gut 2016; 65:963-970

6. Bisschops R, Bessissow T, Dekker E et al Pit pattern analysis with high -definition chromoendoscopy and narrow-band imaging for optical diagnosis of dysplasia in patients with Ulcerative Colitis. Gastrointest Endosc 2017;86:1100-1106

7. Carballal S, Maisterra S, López-Serrano A, et al. Real-life chromoendoscopy for neoplasia detection and characterisation in long-standing IBD. Gut. 2018;67:70-78

8. Iacucci M Kaplan GG , Panaccione R et al A randomized trial comparing high definition colonoscopy alone with high definition dye spraying and electronic virtual chromoendoscopy for detection of colonic neoplastic lesions during IBD surveillance colonoscopy.   Am J Gastroenterol 2018;113:225-234

9.Hata K, Watanabe T, Motoi T, Nagawa H. Pitfalls of pit pattern diagnosis in ulcerative colitis –associated dysplasia .Gastroenterology 2004; 126:374-376

10. Oka S, Tanaka S, Chayama K. Detection of nonpolypoid colorectal neoplasia using magnifying endoscopy in colonic inflammatory bowel disease. Gastrointest Endosc Clin North Am 2014;24:405-417.

11. Rubin DT, Rothe JA, Hetzel JT, et al. Are dysplasia and colorectal cancer endoscopically visible in patients with ulcerative colitis? Gastrointest Endosc 2007;65: 998–1004.

12.Matsumoto T, Iwao Y, Igarashi M, et al. Endoscopic and chromoendoscopic atlas featuring dysplastic lesions in surveillance colonoscopy for patients with long-standing ulcerative colitis. Inflamm Bowel Dis 2008;14:259-64.

13. Sussman DA, Barkin JA, Martin AM, et al. Development of advanced imaging criteria for the endoscopic identification of inflammatory polyps. Clin Translat Gastroenterol 2015;6:e128.

14. Sugimoto S, Naganuma M, Iwao Y  et al. Endoscopic morphologic features of ulcerative colitis-associated dysplasia classified according to the SCENIC consensus statement. Gastrointest Endosc. 2017;85:639-646

> **Comment [SG1]:** REMOVE ?

15. Schlemper RJ, Riddell RH, Kato Y, et al. The Vienna classification of gastrointestinal epithelial neoplasia. Gut. 2000; 47:251-255.

Comment [SG2]: REMOVE ??

16.Fleiss JL, Nee JC, Landis JR. Large sample variance of kappa in the case of different sets of raters. Psychological Bulletin 1979; 86: 974

 17.Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159-174

18. Iacucci M, Hassan C, Fort Gasia M, et al. Serrated adenoma prevalence in inflammatory bowel disease surveillance colonoscopy, and characteristics revealed by chromoendoscopy and virtual chromoendoscopy. Can J Gastroenterol Hepatol 2014;28:589-594.

19. Hewett DG, Kaltenbach T, Sano Y, et al. Validation of a simple classification system for endoscopic diagnosis of small colorectal polyps using narrow-band imaging. Gastroenterology 2012;143:599-607.

20. van Doorn SC,  Hazewinkel Y, East JE et al Polyp Morphology: An Interobserver Evaluation for the Paris Classification Among International Experts Am J Gastroenterol. 2015 ;110:180-187.

21.Raghavendra M, Hewett DG, Rex DK. Differentiating adenomas from hyperplastic colorectal polyps: narrow-band imaging can be learned in 20 minutes. Gastrointest Endosc 2010;72:572–576

22.Rastogi A, Rao DS, Gupta N, et al. Impact of a computer-based teaching module on characterization of diminutive colon polyps by using narrow-band imaging by non-experts in academic and community practice: a video-based study. Gastrointest Endosc 2014;79:390–398

23 Patel SG, Schoenfeld P, Kim HM et al .. Real-Time Characterization of Diminutive Colorectal Polyp Histology Using Narrow-Band Imaging: Implications for the Resect and Discard Strategy. Gastroenterology 2016; 150:406-418

Table 1 : Multivariable analysis of performance characteristics of individual and combined endoscopic criteria to predict neoplastic lesions associated with colitis  histology when used by expert endoscopists.

| Criteria | Post Training | | | | |
|---|---|---|---|---|---|
| | Sensitivity (95%CI) | Specificity (95%) | PPV (95%) | NPV (95%) | Accuracy (95%) |
| **Individual** | | | | | |
| Surface architecture | 85 | 70 | 83 | 74 | 80 |
| Inflammation within lesion | 88 | 13 | 63 | 40 | 61 |
| Paris-morphology nonpolypoid/polypoid | 97 | 12 | 65 | 69 | 66 |
| Vessel architecture | 87 | 36 | 70 | 60 | 69 |
| | | | | | |
| **Two Criteria** | | | | | |
| Surface+Inflammation | 74 | 74 | 83 | 63 | 74 |
| Surface+Paris morphology | 83 | 73 | 83 | 71 | 79 |
| Surface+Vessel | 78 | 76 | 85 | 67 | 78 |
| Inflammation +Paris morphology | 86 | 25 | 66 | 51 | 64 |
| Inflammation+Vessel | 75 | 45 | 71 | 50 | 65 |
| Paris morphology +Vessel | 83 | 45 | 73 | 61 | 70 |
| | | | | | |
| **Three Criteria** | | | | | |
| Surface+Inflammation +Paris morphology | 72 | 77 | 84 | 63 | 74 |
| Surface+Inflammation +Vessel | 67 | 79 | 85 | 58 | 71 |
| Surface + Paris morphology + Vessel | 75 | 78 | 86 | 64 | 77 |
| Inflammation +Paris morphology+Vessel | 73 | 54 | 74 | 53 | 67 |
| | | | | | |
| **All Four** | | | | | |
| Surface+inflammation +Paris | 65 | 81 | 86 | 57 | 71 |

morphology+Vessel

**Table.2 Diagnostic performance for diagnosing colonic Lesions in IBD Pretest Vs Postest Diagnosis Compared to Histology**

| Experts | Pre (95%CI) | Post (95%CI) | P value |
|---|---|---|---|
| Sens | 72(67-76) | 72(67-67) | P = 0.986 |
| Spec | 74(64-82) | 73(61-80) | P= 0.758 |
| Accuracy | 72(68-76) | 72(68-76) | P= 0.718 |
| **Consultants** | | | |
| Sens | 81(75-86) | 90(85-94) | P=0.007 |
| Spec | 68(54-81) | 56(37-74) | P=0.669 |
| Accuracy | 79(73-84) | 86(80-90) | P=0.001 |
| **Registrars/ Trainees** | | | |
| Sens | 80(76-84) | 0.78 82(78-85) | P=0.719 |
| Spec | 54(43-64) | 78(66-87) | P=0.002 |
| Accuracy | 75(71-79) | 81(77-84) | P=1 |
| **Juniors Doctors** | | | |
| Sens | 48(42-54) | 80(75-84) | **P<0.001** |
| Spec | 44(33-57) | 58(43-72) | P=0.111 |
| Accuracy | 48(42-53) | 77(72-81) | **P<0.001** |

| Table 3 . Comparison of performance diagnosis for IBD  colonic lesions histological diagnosis with High vs Low Confidence Post-test | | | |
|---|---|---|---|
| **Experts** | **High Confidence** | **Low Confidence** | **P value** |
| Sens | 72(66-77) | 72(64-79) | P=0.897 |
| Spec | 90(80-96) | 30 (15-49) | P< 0.001 |
| Accuracy | 76(70-80) | 65(57-72) | P= 0.003 |
| | | | |
| **Consultants** | | | |
| Sens | 90(83-95) | 90(82-95) | P= 0.913 |
| Spec | 61(38-80) | 44(13-78) | P=0.403 |
| Accuracy | 85(78-90) | 86(78-92) | P=0.199 |
| | | | |
| **Registrars/ Trainees** | | | |
| Sens | 84(78-89) | 86(81-90) | P=0.357 |
| Spec | 82(64-93) | 70(47-86) | P=0.463 |
| Accuracy | 84(78-88) | 84(80-89) | P=0.618 |
| | | | |
| **Juniors Doctors** | | | |
| Sens | 83(68-93) | 80(75-85) | P=0.344 |
| Spec | 67(35-90) | 54(37-71) | P=0.352 |
| Accuracy | 80(65-90) | 78(72-82) | P=0.457 |

**Table 4. Kappa Coefficient for inter-observer agreement for pre test and post test in the different groups**

| Groups | Pre test kappa (95% CI) | Post test Kappa (95% CI) |
|---|---|---|
| **Consultants** | 0.40 (0.19-0.56) | 0.65 (0.42-0.83) p=0.03 |
| **Registrars/trainees** | 0.30 (0.22-0.44) | 0.41 (0.28-0.58) p=0.13 |
| **Juniors doctors** | 0.15 (0.06-027) | 0.20 (0.09-0.33)  p=0.27 |