An international multicenter real-life prospective study of electronic chromoendoscopy score PICaSSO in ulcerative colitis

Iacucci, Marietta; Smith, Samuel CL; Bazarova, Alina ; Shivaji, Uday; Bhandari, Pradeep; Cannatelli, Rosanna; Daperno, Marco; Ferraz, Jose; Goetz, Martin; Gui, Xianyong; Hayee, Bu ; De Hertogh, Gert ; Lazarev, Mark; Li, Jim ; Nardone, Olga; Parra-Blanco, Adolfo ; Pastorelli, Luca ; Panaccione, Remo; Occhipinti, Vincenzo ; Rath, Timo

DOI:
10.1053/j.gastro.2020.12.024

Document Version
Peer reviewed version

Citation for published version (Harvard):
Iacucci, M, Smith, SCL, Bazarova, A, Shivaji, U, Bhandari, P, Cannatelli, R, Daperno, M, Ferraz, J, Goetz, M, Gui, X, Hayee, B, De Hertogh, G, Lazarev, M, Li, J, Nardone, O, Parra-Blanco, A, Pastorelli, L, Panaccione, R, Occhipinti, V, Rath, T, Tontini, GE, Vieth, M, Villanacci, V, Zardo , D, Bisschops, R, Kiesslich, R & Ghosh, S 2020, 'An international multicenter real-life prospective study of electronic chromoendoscopy score PICaSSO in ulcerative colitis', *Gastroenterology*. https://doi.org/10.1053/j.gastro.2020.12.024

An international multicenter real-life prospective study of electronic chromoendoscopy score PICaSSO in Ulcerative Colitis

Short title: Prospective real-life study of PICaSSO score in UC

Authors: Marietta Iacucci PhD[1,2,3], Samuel CL Smith MRCP[1], Alina Bazarova PhD[1,4], Uday N Shivaji MRCP[1,2], Pradeep Bhandari MD[5], Rosanna Cannatelli MD[1,6], Marco Daperno MD[7], Jose Ferraz PhD[3], Martin Goetz PhD[8], Xianyong Gui MD[9], Bu Hayee PhD[10], Gert De Hertogh PhD[11], Mark Lazarev MD[12], Jim Li MD[13], Olga M Nardone MD[1], Adolfo Parra-Blanco PhD[14], Luca Pastorelli PhD[15], Remo Panaccione MD[3] , Vincenzo Occhipinti MD[15] ,Timo Rath PhD[16], Gian Eugenio Tontini PhD[17], Michael Vieth PhD[18], Vincenzo Villanacci MD[19], Davide Zardo MD[20], Raf Bisschops PhD*[11], Ralf Kiesslich PhD*[21], Subrata Ghosh MD*[1,2,3]

*Contributed equally

Institutions:

1. Institute of Immunology & Immunotherapy, NIHR Wellcome Trust Clinical Research Facilities, University of Birmingham, UK and University Hospitals Birmingham NHS Trust
2. National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre
3. Division of Gastroenterology and Hepatology, University of Calgary, Canada
4. Institute for Biological Physics, University of Cologne, Cologne, Germany
5. Queen Alexandra Hospital, Portsmouth, UK
6. Department of Gastroenterology, Spedali Civili, Brescia, Italy
7. University of Torino, Italy
8. Klinikum Böblingen, Germany
9. University of Washington, Seattle, USA
10. Kings College London, UK
11. University Hospitals Leuven, Belgium
12. Johns Hopkins Hospital, Baltimore, USA
13. Peking Union Medical College Hospital, China
14. University of Nottingham, Nottingham, UK
15. IRCCS Policlinico San Donato, Italy
16. University of Erlangen, Germany
17. Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Department of Pathophysiology and Transplantation, University of Milan, Italy
18. Klinikum Bayreuth, Germany
19. Institute of Pathology Spedali Civili, Brescia, Italy
20. University Hospitals Birmingham NHS Foundation Trust
21. Helios HSK Wiesbaden, Wiesbaden, Germany

Correspondence:

Marietta Iacucci MD, PhD, FASGE, FRCPE

Reader in Gastroenterology

Institute of Immunology and Immunotherapy

University of Birmingham, United Kingdom

Adjunct Clinical Associate Professor of Medicine

University of Calgary, Canada

Address: Institute of Translational Medicine

Heritage Building for Research and Development

University Hospitals Birmingham NHS Foundation Trust

Edgbaston, Birmingham, UK

B15 2TT

Telephone: +44 (0) 121 3718119

Email: m.iacucci@bham.ac.uk

Funding statement:

MI and SG are part-funded by the NIHR Birmingham Biomedical Research Centre The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Disclosures:

None of the other authors have any conflict of interest to declare related to this manuscript.

Author contributions:

MI: Study conception and design, Data acquisition, Analysis and interpretation of data, Drafting of the manuscript, Critical revision of manuscript for important intellectual content

SS: Data acquisition, Analysis and interpretation of data, Drafting of the manuscript, Critical revision of manuscript for important intellectual content

AB: Analysis and interpretation of data, Critical revision of manuscript for important intellectual content, Statistical analysis

US: Data acquisition, Critical revision of manuscript for important intellectual content

PB: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

RC: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

MD: Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

JF: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

MG: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

XG: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

BH: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

GdH: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

ML: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

JL: Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

ON: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

APB: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

LP: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

RP: Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

VO: Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

TR: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

GET: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

MV: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

VV: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

DZ: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

RB: Data acquisition, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

RK: Study conception and design, Analysis and interpretation of data, Critical revision of manuscript for important intellectual content

SG: Study conception and design, Analysis and interpretation of data, Drafting of the manuscript, Critical revision of manuscript for important intellectual content

Abbreviations:

AUROC-Area Under the Receiver Operator Characteristics

ECAP- Extent, Chronicity, Activity, Plus Score

ER- Endoscopic remission

HD WLE- High Definition White Light Endoscopy

HR- Histological remission

IBD- Inflammatory Bowel Disease

ICC – Intra-class Correlation Coefficient

MES- Mayo Endoscopic Score

MH- Mucosal Healing

PICaSSO- Paddington International Virtual ChromoendoScopy ScOre

RHI- Robarts Histopathology Index

ROC- Receiver Operating Characteristics

UC-Ulcerative Colitis

UCEIS- Ulcerative Colitis Endoscopic Index of Severity

VEC- Virtual Electronic Chromoendoscopy

Abstract:

BACKGROUND & AIMS

Endoscopic and Histologic remission are important goals in the treatment of ulcerative colitis (UC). We investigated the correlation of the recently developed Paddington International Virtual ChromoendoScopy ScOre(PICaSSO) and other established endoscopic scores against multiple histological indices and prospectively assessed outcomes.

METHODS:

In this prospective multicenter international study Inflammatory activity was assessed with high definition and virtual chromoendoscopy in the rectum and sigmoid using Mayo Endoscopic Score (MES), UC Endoscopic Index of Severity (UCEIS) and PICaSSO. Targeted biopsies were taken for assessment using Robarts Histological Index (RHI), Nancy Histological index (NHI), ECAP, Geboes and Villanacci. Follow up data was obtained at 6 and 12 months after colonoscopy.

RESULTS:

307 patients were recruited. There was strong correlation between PICaSSO and histology scores, significantly superior to correlation coefficients of MES and UCEIS with histology scores. A PICaSSO score of ≤3 detected histologic remission by RHI (≤3 + absence of neutrophils) with AUROC 0.90 (95% CI 0.86-0.94) and NHI (≤1) AUROC 0.82 (95% CI 0.77-0.87). The inter-observer agreement for PICaSSO was 0.88 (95% CI 0.83-0.92). At 6- and 12-months follow-up PICaSSO score ≤3 predicted better outcomes than PICaSSO >3 {hazard ratio HR 0.19 (0.11-0.33) and 0.22 (0.13-0.34) respectively} as well as PICaSSO 4-8 {HR 0.25 (0.12,0.53) and 0.22 (0.12,0.39) respectively} and similar to histologic remission.

CONCLUSION:

In this first real-life multicenter study, the PICaSSO score correlated strongly with multiple histological indices. Furthemore, PICaSSO score predicted specified clinical outcomes at 6 and 12 months similar to histology. Thus, PICaSSO can be a useful endoscopic tool in the therapeutic management of UC.

Key words: Mucosal Healing; Endoscopic Remission; Histological Remission; Ulcerative Colitis; Virtual Electronic Chromoendoscopy; PICaSSO score

Introduction

Goals of therapy in Ulcerative Colitis (UC) have evolved over time .The International Organization for Study of Inflammatory Bowel Disease (IOIBD) proposed an important treatment paradigm in UC which is termed the "treat to target" strategy, whereby clinicians aim to adjust therapy in order to achieve the target of mucosal healing (MH) [1]. Definition of MH is evolving but currently incorporates endoscopic and histologic remission (ER and HR). There are clear benefits for achieving MH, in terms of maintaining clinical remission, reducing hospitalization and colectomy rates as well as reducing steroid use [2-4]. The Mayo Endoscopic Score (MES)[5] of 0 or 1 is mostly used in clinical studies to define endoscopic improvement . However, relapse rates are higher for MES 1 than MES 0 [6, 7]. Furthermore, amongst patients with MES 0, those who had histologic remission had lower rates of clinical relapse compared with those with histologic activity [8,9]. Currently available advanced imaging using high definition (HD) endoscopy, with virtual electronic chromoendoscopy, (VEC) can enhance mucosal and vascular details[10].

Therefore, there is need for a score that accurately describes ER not just as absence of mucosal changes but with a better correlation to HR. Initial studies have shown that with the current generation of endoscopic equipment with VEC, subtle mucosal and vascular changes may reflect histology [9, 10]. The recently developed, validated and reproduced **P**addington **I**nternational virtual **C**hromoendo**S**copy **S**c**O**re (PICaSSO), is the first to use VEC to assess vascular and mucosal features of healing [11]. Moreover, the PICaSSO score performed better than MES and the Ulcerative Colitis Endoscopic index of Severity (UCEIS)[12] to predict HR in a single center study [11].

Recently, histological assessment of inflammatory activity has been used in clinical trials and been proposed as part of the "treat to target" strategy.[1] A number of histological indices are currently available that include active or chronic changes or both such as Robarts Histopathology Index (RHI)[13], Nancy Histological Index (NHI)[14], Villanacci Simple Score (VSS)[15], Geboes Score[16] and ECAP[9] (Extent, Chronicity, Activity and Plus Score). It has previously been shown that the PICaSSO score using VEC images correlated well with histological scores and had a very good interobserver agreement amongst experts and non-experts [11, 17].

The aim of this first real-life international multicenter study was to prospectively establish the performance and relationship between PICaSSO and other endoscopic scores with several histological scores in UC patients and their association with HR. We also wanted to evaluate whether PICaSSO was predictor of specified clinical outcomes.

Methods

The study was approved by Research Ethics Committee (17/WM/0223) for the UK centers. All international sites gained local research ethics committee approvals in their respective countries. All patients gave informed consent to participate in the study.

Study design:

The study was performed in eleven international centers between September 2016 and November 2019. Adult patients (age ≥18 and ≤80 years) with an established diagnosis of UC were prospectively enrolled when meeting inclusion criteria: a clinical indication for assessment of activity of UC and an established diagnosis of UC for ≥1 year. Exclusion

criteria were inability to provide consent, toxic megacolon, pregnancy or breast feeding, contraindication to biopsies and Boston bowel preparation score <2 [18]. Patients with unclassified colitis, Crohn's colitis, ischemic colitis or infectious colitis were also excluded. In order to study the relationship between PICaSSO and histology, for the first 20 patients recruited, each participating site included quiescent , mild , moderate and severe inflammatory activity based on the clinical partial Mayo score (0-1=quiescent, 2-4=mild, 5-6=moderate, 7-9=severe) at the time of recruitment [5]. Subsequently, to establish the ability of PICaSSO to predict HR, sites were asked to recruit (n=20) mainly mild/quiescent disease (clinical partial Mayo score 0-4), such as patients attending for surveillance purposes or assessment after therapy. This was in recognition of the challenges of MES 0, 1 to relate to HR. All patients who had colonoscopy after consent were included in the study unless they had unsatisfactory bowel preparation.

Study objectives:

The primary objective for this study was to establish the correlation between endoscopic scoring systems for activity defined by PICaSSO [11] (ranging from 0-15), MES [5], UCEIS [12] and histological activity defined by multiple histological scores including RHI [13], NHI [14], GS [16], VSS [15] and ECAP [9].

The secondary objectives were:

(a) to assess ability of PICaSSO score to predict specified clinical outcomes in UC patients at 6 and 12 months after colonoscopy.

(b) to determine the inter-observer agreement of PICaSSO between all raters and compare with other endoscopic scoring systems.

(c) to determine from Receiver Operator Characteristics [ROC] the best thresholds of PICaSSO, MES and UCEIS associated with HR.

The outcome measures were:

(a) Comparison of correlation between endoscopy scores using PICaSSO, MES and UCEIS with the five histology scores RHI, NHI Geboes, Villanacci and ECAP;

(b) Comparison of the prediction of specified clinical outcomes by PICaSSO ≤ 3, MES=0 and UCEIS≤ 1 at 6 months.

Endoscopic assessment:

All endoscopic examinations were performed by gastroenterologists (MI, PB, JF, MG, BH, ML, APB, LP, TR, GT, RB) who were experienced in IBD and optical diagnosis. Prior to the start of the study each center received refresher training materials on i-Scan digital image enhanced endoscopy technology and the PICaSSO score (Figure 1 and video) [11]. Training lasted 1 hour and consisted of examples of annotated images (n=56) and annotated videos (n=4) representative of each mucosal and vascular endoscopic finding.

A polyethylene glycol-based bowel preparation was administered to patients prior to colonoscopy in accordance to each centers protocol. Data collected at the time of the colonoscopy included baseline demographic, duration of disease, extent of colitis and current/previous medication history. Data were recorded on case report forms and transferred to REDCap (The Vanderbilt University, Nashville, Tennessee, USA).

All examinations were performed using HD Pentax (Tokyo, Japan) 7010 processor and colonoscopes of the HiLine series. The mucosa was washed thoroughly prior to the sigmoid colon and rectum being viewed for at least 60 seconds in each four modes: HD, iScan1,

iScan2 and iScan3, which were switched at the push of a button on the endoscope [9]. Standardized settings were used (details in Appendix 1). Endoscopists first assessed the MES [5] and UCEIS [12] with white light- HD followed by PICaSSO score with VEC [11] in the sigmoid and rectum .Two targeted biopsies were taken from the most inflamed area or showing the most representative features of endoscopic remission determined by PICaSSO. The assessment and biopsy locations were captured by video recording for all patients.

Histological assessment:

All rectum and sigmoid biopsy samples were fixed in formalin and then processed at institutional pathology laboratories with routine embedding and staining protocols. The haematoxylin-eosin stained slides of the biopsies were digitized using high-speed slide scanners by participating centers to allow for central reading by the five study pathologists (DZ, MV, VV, GDH and GX), who were blinded to patients' clinical features and endoscopic finding. Each slide was scored using RHI [13], ECAP [9], NHI [14], VSS [15] and GS [16] for comparison with endoscopic grading of inflammatory activity. These histological indices were selected as these represent different combinations of the major histopathologic features of UC, including active inflammation ('activity', i.e., neutrophil infiltration in glandular epithelium and in lamina propria), chronic inflammation (mononuclear cell infiltration in lamina propria) and crypt architectural alteration (chronicity).

For the present study, HR was defined as RHI ≤3 (with absence of neutrophils [19]), ECAP ≤4, NHI ≤1, VSS 0, and GS 0/1. The rectal and sigmoid biopsies were used for correlation with colonoscopic scores from the same site as the biopsies (data of sigmoid biopsies provided in Supplementary material).

Inter-observer agreement:

Participating endoscopists (n=12) and pathologists (n=6) were provided with 40 randomly chosen high quality videos and 50 randomly selected digital histological slides respectively, displaying quiescent and active disease taken from study patients. Each video clip, 60-90 sec in length, was recorded by using the four different modes of HD, iscan1, iscan2 and iscan3. The endoscopists scored each video by using MES, UCEIS and PICaSSO and the pathologists scored each digital slide using RHI, NHI, VSS, GS and ECAP. The data were transferred to case report forms hosted on REDCap (The Vanderbilt University, Nashville, Tennessee, USA).

Follow up and outcomes:

Clinical outcomes at follow up were specified as i) hospitalization as a result of UC relapse ii) colectomy, iii) initiation or changes in medical therapy including steroids, immunomodulators and biologics due to UC relapse. The recorded notes of all patients were reviewed at each center to record these clinical outcome events with telephone calls to clarify any uncertainty, at 6 months and 12 months following colonoscopy procedure in all patients.

Statistical analysis:

The results were exported from REDCap to STATA Version 14 [StataCorp]. Mean ± SD and median ± interquartile range were determined on continuous variables.

Assuming the correlation between PICaSSO- histology and MES- histology to be 0.72 and 0.61 respectively, 302 patients were needed to reach a power of 0.80 with type I error of 0.05 by means of Fisher's Z-transform aiming to address the primary outcome of the study.

On the other hand, 300 patients were needed using 1-sample non-inferiority/superiority test for proportions of specified clinical outcome rates to reach power of 0.80 with type I error of 0.05 when comparing specified clinical outcomes of PICaSSO and MES=0 at 6 months follow-up. We accepted a 10% relapse rate for MES= 0 based on the results of Barreiro-de Acosta et al[7] and 6.5% for PICaSSO score cut-off that was associated with histology score cut-off reflecting HR.

Pearson correlation between various endoscopic scores and histological scores was calculated. Very strong correlation was considered as a value of 0.8-1.0, strong as 0.6-0.79, moderate as 0.40-0.59 and weak as 0.2-0.39. Correlations of endoscopic scores were compared by means of Fisher Z-transformation using R-package (http://www.R-project.org/) cocor. [20]

Inter-observer agreement was assessed using one -way intra class correlation coefficient-ICC (R package irr (http://CRAN.R-project.org/package=irr) . We needed a minimum of 30 videos or digital histological images to reach the power of 0.80 with a type I error of 0.05 to test the hypothesis of the effect size being larger than 0.1 against the alternative of no effect .

We generated 1000 bootstrap sample to compute the differences between PICaSSO and MES/UCEIS ICCs, calculating the 0.01th quantiles of the resulting distributions.

Meaningful misclassification rates were computed splitting each of the endoscopic scores into three categories: remission, mild, moderate/severe (Mayo 0,1,>1; UCEIS≤ 1, 2-4, >4; Picasso≤ 3, 4-8, >8 respectively). For each pair of raters we calculated the difference between their scorings of each video and recorded misclassification if it was greater than zero. For the overall misclassification rate we averaged the obtained values over the

number of videos and the number of pairs. The resulting misclassification rates were statistically compared between each other by computing 1000 values of differences between PICaSSO and UCEIS, PICaSSO and MES misclassification rates, each based on 1000 bootstraps, 0.99th quantiles of these differences.

Receiver Operating Characteristic [ROC] curves were plotted and compared using R package pROC(https://CRAN.R-project.org/package =pROC).[21]Area under the curve [AUROC] was determined to accurately reflect endoscopy scores that predict HR .The point on the curve with the best sensitivity and specificity was determined by Youden's j-statistics [22] and its accuracy was calculated.

Probability of survival without clinical relapse for each of the endoscopic scores was computed using Cox proportional hazard model. Hazard ratios were presented , p values <0.05 were considered statistically significant. To compare the increment in specified clinical outcomes between MES=0 over MES=1 and PICaSSO ≤3 over PICaSSO 4-8 we generated 10000, bootstrap samples from data with MES ≤1 and PICaSS0 ≤8 respectively thereby obtaining two sets of incremental values. We then computed the p values based on the distribution of the differences between each pair of increments.

In order to assess the individual weights of each PICaSSO component we performed a multivariate logistic regression, by means of R-package net [23] (https://CRAN.R-project.org/package=nnet), treating each component of PICaSSO as a binary variable.

Results

Patient demographics:

A total of 307 patients with UC from 11 international centers were prospectively included in the study with corresponding endoscopic and histological scores .A total of 302 patients with complete datasets were included in the final analysis: Demographic features are provided in table 1. At the time of colonoscopy 218 (71.0%) were on 5-ASA, 22 (7.2%) were on corticosteroids, 54 (17.6%) on immunomodulators ,104(33.9%) on biologics and 29 (9.4%) were not on any IBD treatment. None of the patients were on topical therapies or had E1 disease.

The different endoscopy scores correlated with each other. PICaSSO score rectum and sigmoid correlated very strongly with MES rectum 0.81 (95% CI 0.77-0.85); sigmoid 0.82 (95% CI 0.78-0.85) and with UCEIS rectum 0.93 (95% CI 0.91-0.94); sigmoid 0.80 (95% CI 0.76-0.84).

Correlation of endoscopic scores with histological scores:

The correlations between endoscopic and histological scores are shown in figure 2 which presents the heatmap of the distribution of the Pearson correlation coefficients (r) across the endoscopic and histological scores in the rectum. While all the endoscopic scores correlated strongly with histology, the correlation of PICaSSO with all histology scores was significantly better than either MES or UCEIS (Figure 2); this was also the case with the subgroup with the upper range of endoscopy severity scores where PICaSSO correlated significantly better with all histology scores than MES 2/3, UCEIS $\geq$5 (all $p<0.01$; Supplementary Figure 1). The PICaSSO mucosal architectural component very strongly correlated with RHI (r=0.75, 95% CI 0.7-0.8), ECAP (r=0.77, 95% CI 0.72-0.81), NHI (r=0.77, 95% CI 0.72-0.82), GS (r=0.76, 95% CI 0.71-0.81) and VSS (r=0.75, 95% CI 0.69-0.79). The vascular architectural component correlated strongly with RHI (r=0.68, 95% CI 0.62-0.74),

ECAP (r=0.70, 95% CI 0.64-0.76), NHI (r=0.70, 95% CI 0.64-0.76), GS (r=0.68, 95% CI 0.62-0.74) and VSS (r=0.69, 95% CI 0.62-0.74) respectively. The correlation between PICaSSO rectum and sigmoid was 0.66 (95% CI 0.59, 072). PICaSSO sigmoid correlated better with all histology scores than MES, but for UCEIS only for the upper end of the range. (supplementary Figure 2).

Relationship between histological remission and endoscopic scores

95% and 74% of patients with HR by RHI in rectum had PICaSSO $\leq$3 and MES =0 (PICaSSO vs MES p<0.00001). 94% and 69% of patients with HR by RHI in sigmoid had PICaSSO $\leq$3 and MES=0 (PICaSSO vs MES p<0.00001). 97% and 77% of patients with HR by NHI in rectum had PICaSSO $\leq$3 and MES =0 (PICaSSO vs MES p<0.0001). 94% and 73% of patients with HR by NHI in sigmoid had PICaSSO $\leq$3 and MES=0 (PICaSSO vs MES p<0.0001). The false negative rate of PICaSSO score for HR by RHI and NHI was significantly lower than the false negative rate for HR determined by MES.

The Receiver Operating Characteristics (ROC)

The Receiver Operating Characteristics (ROC) of the endoscopic scores in the rectum associated with HR using RHI, ECAP, NHI, GS, VSS are presented in Figure 3. The ability of endoscopic scores indicating HR in terms of sensitivity, specificity and accuracy are presented in table 2 (sigmoid data presented in Supplementary Figure 3 and Supplementary Table 1).

Relationship between RHI $\leqslant$3 and endoscopic scores (Figure 3a):

Defining HR as RHI ≤3 in the absence of neutrophils in the epithelium and lamina propria, a PICaSSO total score of ≤3 was associated with HR with an AUROC of 0.90 (95% CI 0.86-0.94),

UCEIS score of ≤1 was associated with HR with an AUROC of 0.89 (95% CI 0.85-0.94) and a MES of 0 was associated with HR with an AUROC of 0.87 (95% CI 0.83-0.91).

Relationship between ECAP ⩽4 and endoscopic scores (Figure 3b):

The PICaSSO total score ≤3 score was associated with HR defined as ECAP ≤4 with an AUROC of 0.77 (95% CI 0.72-0.82), UCEIS ≤1 was associated with HR with an AUROC 0.74 (95% CI 0.69-0.80) and MES of 0 was associated with HR with an AUROC 0.74 (95% CI 0.69-0.79).

Relationship between NHI ⩽1 and endoscopic scores (Figure 3c):

HR defined as a NHI of ≤1 was associated with PICaSSO total score ≤3 score with an AUROC of 0.82 (95% CI 0.77-0.87), UCEIS ≤1 was associated with HR with an AUROC 0.79 (95% CI 0.73-0.84) and MES of 0 was associated with HR with an AUROC 0.81 (95% CI 0.77-0.86).

Relationship between GS ⩽1 and endoscopic scores (Figure 3d):

The PICaSSO total score ≤3 score was associated with HR defined by a GS ≤1 with an AUROC 0.71 (95% CI 0.66-0.76), the UCEIS ≤1 was associated with HR with AUROC 0.76 (95% CI 0.71-0.81) and MES of 0 was associated with HR with an AUROC 0.68 (95% CI 0.63-0.73).

Relationship of VSS of 0 with endoscopic scores (Figure 3e):

The PICaSSO total score ≤3 score was associated with HR as defined by VSS of 0, with an AUROC 0.88 (95% CI 0.84-0.93), the UCEIS ≤1 was associated with HR with an AUROC 0.86 (95% CI 0.81-0.91) and MES of 0 was associated with HR with AUROC 0.87 (95% CI 0.82-0.91).

Relationship between RHI and Nancy with PICaSSO mucosal and vascular architecture

A PICaSSO mucosal score of ≤1 and PICaSSO vascular score of 0 predicted HR as defined by RHI of ≤3 in the absence of neutrophils in the epithelium and lamina propria, both with an AUROC 0.88 (95% CI 0.83-0.92). A PICaSSO mucosal score of ≤1 and PICaSSO vascular score of 0 was associated with HR as defined by NHI of ≤1 with an AUROC 0.80 (95% CI 0.76-0.86) and 0.80 (95% CI 0.75-0.85) respectively.

Inter-observer agreement of PICaSSO, UCEIS and MES scores:

The intra-class correlation coefficient (ICC) for the MES, UCEIS total and PICaSSO total score by the twelve expert endoscopists using high quality videos are shown in Table 3. The one-way ICC between raters was 0.88 (95% CI 0.83-0.92) for the overall PICaSSO score, as compared with ICC MES of 0.82 (95% CI 0.74-0.88) and ICC UCEIS total score of 0.84 (95% CI 0.77-0.87). The PICaSSO vascular component had a good agreement ICC=0.74 (95% CI 0.64-0.83) whereas the vascular component of the UCEIS had a moderate agreement ICC=0.55 (95% CI 0.44-0.64).

PICaSSO ICC was significantly ($p<0.01$) higher than ICC of either MES or UCEIS. In addition, the PICaSSO misclassification rate was significantly lower than that of MES and UCEIS {PICaSSO 0.12, UCEIS 0.24, MES 0.29. ($p<0.01$ PICaSSO Vs MES, UCEIS)}.

Inter-observer agreement of the five histology scores is shown in table 3 .

Follow up results of specified clinical outcomes :

A total of 289 and 270 patients completed 6 and 12 months follow up respectively after colonoscopy. A total of 66 (22.8%) and 78 (28.9%) patients had a specified clinical outcome events at 6 and 12 months respectively. Figure 4 and Figure 5 presents all the Cox

proportional hazards curves for the PICaSSO, UCEIS and MES (rectum) at 6 and 12 months, respectively.

There was a survival benefit without specified clinical outcomes for PICaSSO score of ≤3 vs >3, UCEIS ≤1 vs. >1 and MES 0 vs. MES ≥1 .MES 0 had a significantly better outcome over MES 1 at both 6 and 12 months. In addition, analysis of PICaSSO ≤3 vs PICaSSO >3 ≤8 demonstrated significant survival benefit without specified clinical outcomes at 6 and 12 months .The sigmoid data is shown in Supplementary Figure 4. The endoscopy scores had similar specified clinical outcomes at 6 and 12 months.

A further sub-analysis to demonstrate the ability of the PICaSSO mucosal and PICaSSO vascular architectural scores separately in the rectum to predict specified clinical outcomes was performed (Supplementary figure 5). There was a significantly increased survival benefit without specified clinical outcomes with a vascular score of 0 vs. ≥1 or mucosal score of <2 vs. ≥2 at 6 and 12 months .The sigmoid data is shown in Supplementary Figure 6.

Based on the bootstrap samples the increment in specified clinical outcomes (event rates) of MES=1 over MES=0 was significantly greater than increment of events of PICaSSO 4 – 8 over PICaSSO ≤3 at 6 months follow-up ($p<0.05$ for both rectum and sigmoid). The difference between MES 0 and 1 is therefore significantly bigger than PICaSSO ≤3 and 4-8.

In patients with PICaSSO ≤3 (rectum) the specified clinical outcomes at 6 and 12 months were 12.7% and 16.8% respectively . This was not significantly different from PICaSSO =0

(10.8% and 14.8% at 6 and 12 months). In patients with RHI≤3 and no neutrophils the specified clinical outcomes events at 6 and 12 months were 11.7% and 15.9% respectively . NHI ≤1 rates for specified clinical outcomes events at 6 and 12 months were similar (12.2%

and 17.1%). Therefore, PICaSSO and histology specified clinical outcome rates after remission were similar.

We also analyzed the weightage of each component of the PICaSSO score (mucosal and vascular) by means of multivariate logistic regression in order to predict specified clinical outcome at follow-up, which is shown in supplementary table 2. We note that mucosal component I-B (patchy micro-erosion) and vascular component II-B (intra-mucosal bleeding, crowded or tortuous superficial vessels with dilatation) contributed relatively less to the regression score.

Discussion

An endoscopic scoring system that accurately encompasses the elements of MH ,namely ER and HR is needed. We have reported, for the first time a large prospective multi-center, real-life study of the PICaSSO score in UC for assessment of ER and HR. The biopsies were taken at the same site of endoscopic assessment and scored by pathologists in a blinded fashion. The results show that PICaSSO score correlate strongly with multiple histological indices, and the correlation coefficients were statistically superior to both MES (rectum and sigmoid) and UCEIS (rectum only); this also applied to the upper end of the endoscopy disease spectrum for both MES and UCEIS.

In this study the PICaSSO score had a rating of almost perfect inter-observer agreement, which is consistent with our previous study [11]. The ICC of PICaSSO was significantly better than MES or UCEIS, with less meaningful misclassification rate. We have previously shown

that the PICaSSO score could be reproduced in less experienced and trainee endoscopists after a short training module [17]

Commonly used endoscopic scores of inflammatory activity in UC are MES and UCEIS; however these scores were developed with previous generation of endoscopes and have limitations, nor least that these do not truly reflect ER [9] and HR[24-26]. HR my predict of favorable patient outcomes [27-29], and we have demonstrated that the PICaSSO score of ≤3 (not only describes endoscopically quiescent disease but is also associated with HR with a high degree of accuracy and low false negative rate. This confirms our recent findings with a single operator using multiple endoscopic platforms [10], with clinical and fecal calprotectin endpoints[30] reinforcing the concept that endoscopic assessment is getting closer to histology using current generation endoscopes. VEC is a technology that is widely available at a push of a button ,routinely used for better characterization of GI cancers and do not require additional time [30,35]

Several studies have shown that even MES 0 was imperfect for providing a definition of MH as histological inflammatory activity is frequently present [9,10, 26]. However, our results in this study showed that the MES =0 and UCEIS ≤1 ,also detected HR. This may be the consequence of assessment by HD, experienced operators and training modules prior to start of the study. [9, 10,31 32] UCEIS in a previous study reported moderate inter-observer agreement [33] and correlated strongly with the NHI and RHI, but this study was single center, single endoscopist and pathologist design [34].

For an endoscopic score to be useful clinically, a cut-off to indicate likelihood of long-term outcome of remission is desirable. We have shown that there is a significant favorable outcome of PICaSSO score ≤3 over higher scores at both 6 and 12 months. In addition, we

have demonstrated an outcome benefit of patients with PICaSSO ≤3 vs. PICaSSO 4-8, which allows a greater bandwidth when defining ER; The mucosal and the vascular components of the PICaSSO score were independently able to predict outcomes on follow-up.

MES and UCEIS also performed very well in our study as these were also assessed using HD imaging and scored by experienced operators using targeted biopsies .

The main strengths of the PICaSSO score are the following: strong correlation with all histology scores (better than MES and UCEIS) , accurate correlation with HR, excellent reliability as demonstrated by its excellent ICC with lowest misclassification rate, lowest false negative rate of PICaSSO≤3 for HR and it can predict accurately outcomes. Targeted biopsies contributed to strong endoscopy-histology correlations.[10, 30]. Specified clinical outcome rates were similar for patients in remission by PICaSSO and by RHI or NHI. The endoscopy scores performed similarly for outcomes. Of note, specified clinical outcomes rates had a larger increase between MES 0 and 1 compared with PICaSSO ≤3 and 4-8, a limitation of a 4-point MES and the challenge of considering MES=0 and 1 as endoscopic improvement.

A future use of the PICaSSO score could be in clinical trials, where precise assessment of inflammation and ER from endoscopic videos in central read-out [36] is crucial in determining response to novel treatments. Unlike MES which is a composite score with a number of descriptors for each score item, PICaSSO mucosal and vascular scoring items each describe a single feature, and hence may be suitable for future development of artificial intelligence . A summary of the characteristics of PICaSSO against other endoscopic scores is presented in Supplementary Table 3.

A limitation of the study is the participation of endoscopists experienced in optical diagnosis. While ,this could be seen as a strength as it allows a better understanding of the promise of PICaSSO and VEC, it cannot be assumed the same levels of performance would be replicated amongst non-experts and not accustomed to VEC in a real-life setting. However, we have previously shown that by using a short training module on PICaSSO, the level of performance can be reproduced even in non-expert endoscopists and trainees [17]. PICaSSO is compatible with other VEC platforms such as NBI and BLI as shown by us [17, 30]. All raters underwent refresher training for PICaSSO using a structured and previously validated training module and this may have improved performance also for MES and UCEIS, though MES, UCEIS and PICaSSO were scored sequentially.[17] We did not specify to assess fecal calprotectin in this study but we had shown correlation with PICaSSO previously using full colonoscopy[30] We also did not follow-up patients using patient reported outcomes similar to others study [7, 37] as symptoms do not relate well to endoscopy and histology. In case of 6- (and 12) months follow-up the required sample size was not reached and even though initially 307 patients were recruited , only 289 and 270 completed the respective follow-up periods. The number of PICaSSO ≤3 and MES=0 patients who completed follow-ups was even lower.

In conclusion, we present a large, real life, prospective multicenter study that shows the first VEC PICaSSO is able to detect ER and HR accurately, and has a strong correlation across the range of endoscopy scores with histological activity which is better than the MES and UCEIS scores. PICaSSO should be considered in clinical practice and clinical trials to grade endoscopic inflammatory activity and define MH and further prospective studies in the context of specific therapies are necessary to assess responsiveness and outcomes.

Acknowledgements: We would like to thank our excellent research nurses Christine Bossa and Rosemary Matonhodze and our research coordinators Hollie Caulfield, Suzanne Lockyer, Gail Heritage and Helen Duffy and pathologist Dr Rahul Hejmadi who contributed constructively to several components of the study.

Figure legends:

Figure 1: Composite figure explaining all the endoscopic features of PICaSSO and corresponding histology images

Figure 2. Heatmap demonstrating Pearson correlation between endoscopic and histological scores in the rectum for all patients. Red indicates a weaker correlation and white indicates stronger correlation

Figure 3. Receiver operator curves (ROC) for the PICaSSO, UCEIS and MES in predicting histological healing a) RHI ≤3 no neutrophils, b) ECAP ≤4, c) Nancy ≤1, d) Geboes ≤1 and e) Villanacci 0

Figure 4. Cox proportional hazard curves in predicting likelihood of specified clinical outcomes at 6 months of a) PICaSSO ≤3 (blue) vs. >3 (red) b) UCEIS ≤1 (blue) vs. >1 (red) C) MES 0 vs. ≥1 d) MES 0 (blue) vs. MES 1 (red) and e) PICaSSO ≤3 (blue) vs. 4-8 (red) calculated in the rectum

Figure 5 .Cox proportional hazard curves in predicting likelihood of specified clinical outcomes at 12 months of a) PICaSSO ≤3 (blue) vs. >3 (red) b) UCEIS ≤1 (blue) vs. >1 (red) C) MES 0 vs. ≥1 d) MES 0 (blue) vs. MES 1 (red) and e) PICaSSO ≤3 (blue) vs. 4-≤8 (red) calculated in the rectum

Supplementary data:

Supplementary figure 1. Heatmap demonstrating Pearson correlation between endoscopic and histological scores in patients with moderate/severe endoscopic activity in the rectum (left heatmap) and in the sigmoid (right heatmap). Red indicates a weak correlation and white indicates strong correlation

Supplementary figure 2. Heatmap demonstrating Pearson correlation between endoscopic and histological scores in the sigmoid colon for all patients. Red indicates a weak correlation and white indicates strong correlation

Supplementary Figure 3. Receiver operator curves (ROC) for the PICaSSO, UCEIS and MES in predicting histological healing in the sigmoid colon a) RHI ⩽3 no neutrophils, b) ECAP ⩽4, c) Nancy ⩽1, d) Geboes ⩽1 and e) Villanacci 0

Supplementary figure 4. Cox proportional hazard curves in predicting likelihood of specified clinical outcomes at 6 months a) PICaSSO ≤3 (blue) vs. >3 (red) b) UCEIS ≤1 (blue) vs. >1 (red) c) PICaSSO ≤3 (blue) vs. 4--8 (red); and 12 months d) PICaSSO ≤3 (blue) vs. >3 (red) e) UCEIS ≤1 (blue) vs. >1 (red) f) PICaSSO ≤3 (blue) vs. 4-8 (red) calculated in the sigmoid colon

Supplementary figure 5. Cox proportional hazard curves of PICaSSO mucosal score and PICaSSO vascular architectural scores predicting specified clinical outcomes at 6 (A and B) and 12 months (C and D) in the rectum

Supplementary figure 6. Cox proportional hazard curves of PICaSSO mucosal score and PICaSSO vascular architectural scores predicting clinical outcomes at 6 (A and B) and 12 months (C and D) in the sigmoid colon

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. Shah SC, Colombel JF, Sands BE, et al. Systematic review with meta-analysis: mucosal healing is associated with improved long-term outcomes in Crohn's disease. Aliment Pharmacol Ther 2016;43:317-33.
3. Ferrante M, Vermeire S, Fidder H, et al. Long-term outcome after infliximab for refractory ulcerative colitis. J Crohns Colitis 2008;2:219-25.
4. 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.
5. 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.
6. Boal Carvalho P, Dias de Castro F, Rosa B, et al. Mucosal Healing in Ulcerative Colitis--When Zero is Better. J Crohns Colitis 2016;10:20-5.
7. 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.
8. Yoon H, Jangi S, Dulai PS, et al. Incremental Benefit of Achieving Endoscopic and Histologic Remission in Patients With Ulcerative Colitis: A Systematic Review and Meta-Analysis. Gastroenterology 2020.
9. Iacucci M, Fort Gasia M, Hassan C, et al. Complete mucosal healing defined by endoscopic Mayo subscore still demonstrates abnormalities by novel high definition colonoscopy and refined histological gradings. Endoscopy 2015;47:726-34.
10. Iacucci M, Cannatelli R, Gui XS, et al. Assessment of endoscopic healing by using advanced technologies reflects histologic healing in ulcerative colitis. J Crohns Colitis 2020.
11. Iacucci M, Daperno M, Lazarev M, et al. Development and reliability of the new endoscopic virtual chromoendoscopy score: the PICaSSO (Paddington International Virtual ChromoendoScopy ScOre) in ulcerative colitis. Gastrointest Endosc 2017;86:1118-1127 e5.
12. Travis SP, Schnell D, Krzeski P, et al. Developing an instrument to assess the endoscopic severity of ulcerative colitis: the Ulcerative Colitis Endoscopic Index of Severity (UCEIS). Gut 2012;61:535-42.
13. Mosli MH, Feagan BG, Zou G, et al. Development and validation of a histological index for UC. Gut 2017;66:50-58.
14. Marchal-Bressenot A, Salleron J, Boulagnon-Rombi C, et al. Development and validation of the Nancy histological index for UC. Gut 2017;66:43-49.
15. Villanacci V, Antonelli E, Lanzarotto F, et al. Usefulness of Different Pathological Scores to Assess Healing of the Mucosa in Inflammatory Bowel Diseases: A Real Life Study. Sci Rep 2017;7:6839.
16. Geboes K, Riddell R, Ost A, et al. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. Gut 2000;47:404-9.
17. Trivedi PJ, Kiesslich R, Hodson J, et al. The Paddington International Virtual Chromoendoscopy Score in ulcerative colitis exhibits very good inter-rater agreement after computerized module training: a multicenter study across academic and community practice (with video). Gastrointest Endosc 2018;88:95-106 e2.
18. Calderwood AH, Jacobson BC. Comprehensive validation of the Boston Bowel Preparation Scale. Gastrointest Endosc 2010;72:686-92.
19. Pai RK, Khanna R, D'Haens GR, et al. Definitions of response and remission for the Robarts Histopathology Index. Gut 2019;68:2101-2102.

20. Diedenhofen B, Musch J. cocor: a comprehensive solution for the statistical comparison of correlations. PLoS One 2015;10:e0121945.
21. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.
22. Youden WJ. Index for rating diagnostic tests. Cancer 1950;3:32-5.
23. Venables WN, Ripley BD, Venables WNMaswSP. Modern applied statistics with S. New York: Springer, 2002.
24. Kanazawa M, Takahashi F, Tominaga K, et al. Relationship between endoscopic mucosal healing and histologic inflammation during remission maintenance phase in ulcerative colitis: a retrospective study. Endosc Int Open 2019;7:E568-e575.
25. Bryant RV, Winer S, Travis SP, et al. 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.
26. 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.
27. Christensen B, Hanauer SB, Erlich J, et al. Histologic Normalization Occurs in Ulcerative Colitis and Is Associated With Improved Clinical Outcomes. Clin Gastroenterol Hepatol 2017;15:1557-1564.e1.
28. Calafat M, Lobaton T, Hernandez-Gallego A, et al. Acute histological inflammatory activity is associated with clinical relapse in patients with ulcerative colitis in clinical and endoscopic remission. Dig Liver Dis 2017;49:1327-1331.
29. Park S, Abdi T, Gentry M, et al. Histological Disease Activity as a Predictor of Clinical Relapse Among Patients With Ulcerative Colitis: Systematic Review and Meta-Analysis. Am J Gastroenterol 2016;111:1692-1701.
30. Cannatelli R, Bazarova A, Zardo D, et al. Fecal Calprotectin Thresholds to Predict Endoscopic Remission Using Advanced Optical Enhancement Techniques and Histological Remission in IBD Patients. Inflammatory Bowel Diseases 2020.
31. Limdi JK, Picco M, Farraye FA. A review of endoscopic scoring systems and their importance in a treat-to-target approach in inflammatory bowel disease (with videos). Gastrointest Endosc 2020;91:733-745.
32. Iacucci M, Kiesslich R, Gui X, et al. Beyond white light: optical enhancement in conjunction with magnification colonoscopy for the assessment of mucosal healing in ulcerative colitis. Endoscopy 2017;49:553-559.
33. Travis SP, Schnell D, Krzeski P, et al. Reliability and initial validation of the ulcerative colitis endoscopic index of severity. Gastroenterology 2013;145:987-95.
34. Irani NR, Wang LM, Collins GS, et al. Correlation Between Endoscopic and Histological Activity in Ulcerative Colitis Using Validated Indices. J Crohns Colitis 2018;12:1151-1157.
35. Iacucci M, Furfaro F, Matsumoto T, et al. Advanced endoscopic techniques in the assessment of inflammatory bowel disease: New technology, new era. Gut 2019;68:562-572.
36. 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.
37. 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.