Long-term follow-up of the red density pilot trial: a basis for long-term prediction of sustained clinical remission in ulcerative colitis?

OPEN ACCESS

Authors
Pieter Sinonquel[1,2], Peter Bossuyt[2,3], João Pedro Guedelha Sabino[1,2], Bram Verstockt[1,2], Takao Makino[4], Severine Vermeire[1,2], Raf Bisschops[1,2]

Institutions
1   Department of Translational Research in Gastrointestinal Disorders (TARGID), KU Leuven, Leuven, Belgium
2   Department of Gastroenterology and Hepatology, KU Leuven University Hospitals Leuven Gasthuisberg Campus, Leuven, Belgium
3   Gastroenterology, Imelda General Hospital, Bonheiden, Bonheiden, Belgium
4   Pentax Life Care Division, Hoya Corporation Showanomori Technical Center Pentax Life Care Division, Akishima, Japan

Corresponding author
Dr. Pieter Sinonquel, KU Leuven, Department of Translational Research in Gastrointestinal Disorders (TARGID), Leuven, Belgium
pieter.sinonquel@uzleuven.be

Additional material is available at https://doi.org/10.1055/a-2153-7210

ABSTRACT

Red density (RD) technology is an automated operator-independent endoscopic scoring system for disease activity in ulcerative colitis (UC). In this retrospective analysis we aimed to assess the predictive value of the RD sore for sustained clinical remission. All 39 patients from the RD pilot trial were evaluated for clinical outcome in a 5-year period. The highest RD score was considered for Receiver operating characteristic (ROC) analysis to determine the cut-off of the RD for the composite endpoint of treatment failure (defined as mortality, colectomy, hospitalizations, flares and UC therapy changes). Statistical significance was considered $P < 0.05$. Reassessment of the RD score was possible in 36 patients. The composite endpoint was reached in 17 of 39 patients. ROC analysis for clinical remission showed a RD cut-off of 65, area under the ROC was 0.68, sensitivity of 0.71, and a specificity of 0.63. A RD score of ≥ 65 demonstrated a statistically non-significant increase in composite endpoint (hazard ratio 0.49 (95% confidence interval 0.1871-1.280); $P = 0.1453$). In conclusion, the RD score may be an independent predictor of clinical remission in patients with UC for the disease course up to 5 years, but results of the ongoing PROCEED-UC trial are to awaited for definite conclusions.

Introduction

Ulcerative colitis (UC) is a chronic multifactorial inflammatory disease of the colon (IBD) characterized by a relapsing disease course. Therapeutic algorithms like treat to target or exit strategies lean on the mucosal healing paradigm because it impacts long-term clinical outcome [1, 2, 3]. Recently, the definition of mucosal healing evolved to include both endoscopic and histo-

logic improvement. However, assessment of endoscopic healing is challenging, because it suffers from important inter-rater and intra-rater variability [4]. Based on expert opinion, the leading definition of endoscopic remission is an Ulcerative Colitis Endoscopic Index of Severity (UCEIS) or Mayo Endoscopic Score (MES) of 0 and for endoscopic response, a decrease in MES ≥ 1 grade or a decrease in UCEIS ≥ 2 points [5].

Nevertheless, there still remains uncertainty about what degree of remission is needed for a favorable long-term outcome [6, 7]. Moreover, the depth of remission varies based on the type of outcome one wants to achieve. Hence, histological disease activity may be a better predictor of disease relapse than endoscopic disease activity. Objective scoring systems accounting for minimal mucosal inflammatory changes are lacking, necessitating biopsy sampling and human pathological interpretation and scoring [8].

The novel red density (RD) technology (Pentax Medical, Tokyo, Japan) is the first objective operator-independent endoscopic scoring system for disease activity in UC. Previous research demonstrated a good correlation with both endoscopic and histological disease activity with a RD cut-off for histological remission of 60 [9]. In this retrospective analysis, we aimed to assess the predictive value of the RD sore for sustained clinical remission over a 5-year period.

Materials and methods

Study population

From March 2017 to August 2017, consecutive patients with UC were screened for enrollment in our pilot and development trials. In total, 39 patients with UC and six healthy controls were included and demographics have been described previously [9]. A 5-year follow-up period was predetermined. The Ethics Committee UZ Leuven approved the study (s63012). All patients provided informed consent prior to enrollment.

Red density technology

The RD technology (Pentax Medical, Tokyo, Japan) is based on an algorithm for automatic computer-aided assessment of the redness on a pixel level. The algorithm was prospectively constructed in three phases. The first construction phase consisted of testing the feasibility of the RD algorithm with consecutive optimization for correlation with endoscopic (MES and UCEIS) and histologic scoring systems (Geboes score and Robarts' Histological Index). The second phase tested the operating properties of the RD score developed in Phase 1 in patients with UC that needed treatment escalation for a disease flare (23 patients and 6 health controls). In the third phase, the constructed algorithm from Phase 1 was validated in a validation cohort (16 patients). A detailed overview of the three-phase construction of the RD algorithm was previously described [9]. Initial imaging, as described previously, was reassessed by the Product Development Department of Pentax Medical for the final established RD score (▶Fig. 1). The segment with the highest RD score was considered as relevant for statistical analysis.
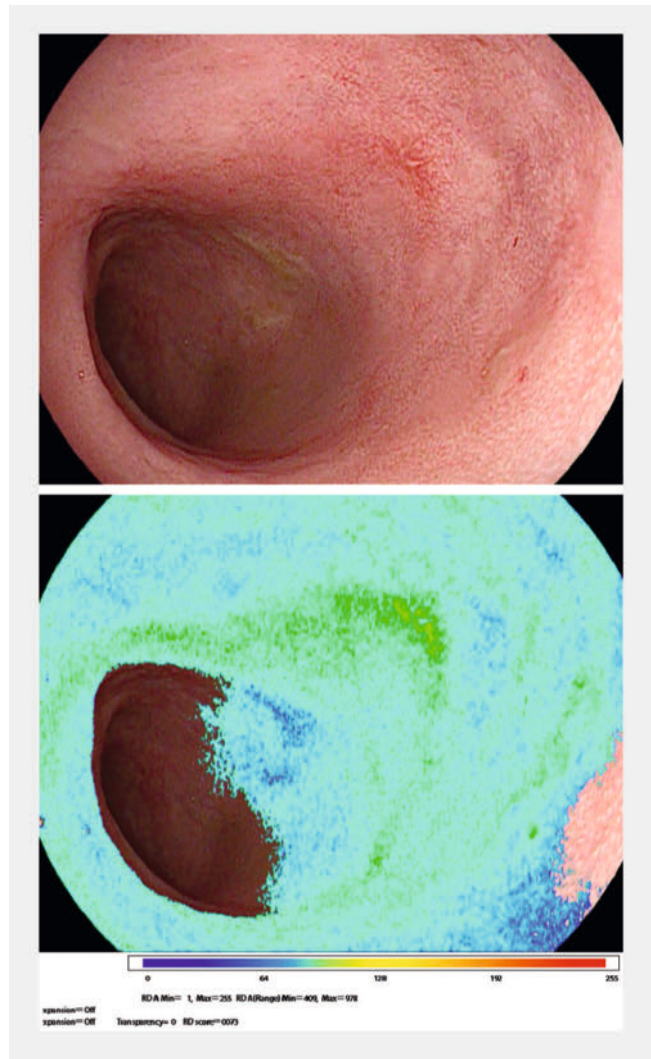
▶ **Fig. 1** Endoscopic image of MAYO 2 colitis in the sigmoid during white light endoscopy, and the corresponding representation of the RD score indicating a score of 73 compatible with non-remission.

Endpoints

The composite endpoint of treatment failure for this 5-year follow-up was defined as change in clinical outcome other than remission. The studied clinical outcomes were: (1) mortality; (2) need for colectomy due to refractory disease; (3) disease flares; (4) IBD-related hospitalizations; and (5) need for change in treatment.

The primary endpoint was defined as the correlation between the composite endpoint and sustained clinical remission based on the optimal RD score.

Secondary outcomes were correlation of the optimal RD score with the individual components of the composite endpoint.

Statistical analysis

Statistical analysis was performed using Graphpad Prism version 9.4.1 (Graphpad Software Inc). Continuous variables with a non-normal distribution were described as median (interquar-

tile range. Categorical variables were described as percentages and the χ^2 tests (or Fisher's exact test, if applicable) were used. For analysis of the optimal cut-off, receiver operating characteristics curve (ROC) analysis was used. Correlation between the primary endpoint and the optimal RD score cut-off was represented in a Kaplan-Meier curve. Linear regression models were used for the calculation of the correlation between RD score cut-off and the different clinical outcome measures described as secondary endpoints. Statistical significance level was set at two-sided $P < 0.05$.
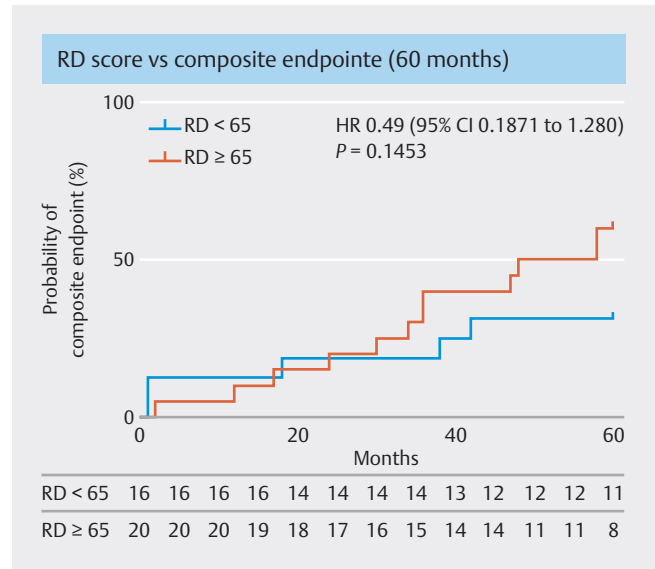
Results

Prediction of composite endpoint of treatment failure

A total of 39 patients were included for retrospective analysis of reaching the composite endpoint within a follow-up period of 5 years ranging from March 2017 to August 2022. Three patients were lost to follow-up, resulting in 36 (92.3%) eligible for analysis. Seventeen patients (47.2%) reached the composite endpoint, in four of whom it was due to colectomy, 13 due to at least one disease flare with change of therapy and/or hospitalization, and no patients died during the follow-up period.

Based on ROC-analysis, the graphical optimal cut-off for prediction of long-term clinical remission was a RD score of 65 (Supplementary Fig. 1). The corresponding area under the ROC was 0.68 with a sensitivity of 0.71 and a specificity of 0.63 with a likelihood ratio of 1.916, resulting in a positive predictive value (PPV) of 0.65 and a negative predictive value of 0.68. There were 16 patients (44.4%) with a RD score of < 65 and 20 (55.6%) with a RD score of ≥ 65 (▶ Table 1). Of the 16 patients with a RD score < 65, only five (31.3%) reached the composite treatment failure endpoint in a median of 18 months (39), due to one colectomy and four disease flares. In the group of patients with a RD score ≥ 65, the composite endpoint was reached in 12 patients (60.0%) after a median of 35 months (27). A RD score of ≥ 65 demonstrated a statistically non-significant increase in composite endpoint (hazard ratio 0.49 [95% confidence interval 0.1871–1.280]; $P = 0.1453$) (▶ Fig. 2).

Prediction of individual parameters for treatment failure

For assessment of the predictive value of the RD score, a cut-off of 65 to the individual parameters of treatment failure was defined as: (1) mortality; (2) need for (partial) colectomy; (3) number of disease flares; (4) number of hospitalizations; and (5) need for step-up of therapy or treatment escalation, and a linear regression model was drawn for each of these. Among the 16 patients with a RD score < 65, one patient (6.3%) underwent colectomy, one (6.3%) needed treatment escalation, and three (18.8%) had one or more flares, of whom, one patient required hospitalization. In the RD score ≥ 65 group, two (10%) had a colectomy and 10 (50.0%) had one or more flares and two needed hospitalization (▶ Table 1). Linear regression modeling showed low correlation with colectomy (r = 0.07, $P = 0.6455$), number of flares (r = 0.05, $P = 0.7292$), number of hospitalizations (r = 0.15, $P = 0.3880$), and number of therapy

RD score vs composite endpointe (60 months)

HR 0.49 (95% CI 0.1871 to 1.280)
$P = 0.1453$

RD < 65	16	16	16	16	14	14	14	14	13	12	12	12	11
RD ≥ 65	20	20	20	19	18	17	16	15	14	14	11	11	8

▶ **Fig. 2** Red density score cut-off of 65 for prediction of reaching the composite endpoint for treatment failure (defined as mortality, colectomy, hospitalizations, flares and UC therapy changes) over 60 months.

changes (r = 0.14, $P = 0.4318$). No regression model was drawn for mortality because no patients died during this 5-year follow-up period.

Discussion

In this cohort study, automated analysis of endoscopic disease activity in patients with UC using the RD index showed positive trends for prediction of meaningful long-term outcomes in UC.

The RD score (Pentax Medical, Tokyo, Japan) has already been demonstrated to correlate well with both endoscopic and histologic scoring systems for assessment of disease activity [9]. However, the correct assessment of mucosal healing remains difficult with great inter-rater and intra-rater variability, even in expert hands [4]. A deep learning model based on convolutional neural network methodology by the group of Takenaka et al. was shown to significantly impact clinical outcome (need for hospitalization, colectomy, steroid-use and relapse) by the automated prediction of mucosal healing [10]. Such automated systems could identify deep mucosal healing without a mucosal biopsy specimen, offering an objective, consistent, and real-time mucosal evaluation with predictive value to the patient.

Because the selection of the "optimal" cut-off is in some way arbitrary, we emphasize that another, more accurate cut-off for prediction of long-term clinical remission is possible. In the assumption of requiring an RD score with a rather high specificity and PPV, the arbitrary cut-off of 100 has been investigated and showed an increase in specificity, although with very low sensitivity and no incremental PPV. Therefore, before drawing more definite conclusions about prediction of long-term clinical outcome, further prospective data and optimization of ROC-analysis is needed.

▶ **Table 1** Number of patients reaching composite endpoint or individual clinical outcome.

	RD score < 65	RD score ≥ 65	
Patients, n (%)	16 (44.4)	20 (55.6)	
Composite endpoint, n (%)	5 (31.3)	12 (60.0)	HR 0.49 (95% CI 0.1871–1.280); $P = 0.1453$
Median time, n (IQR)	18 (39)	35 (27)	
Mortality, n (%)	0 (0.0)	0 (0.0)	
Colectomy, n (%)	1 (6.3)	2 (10.0)	$r = 0.07$
Partial, n (%)	0 (0.0)	1 (50.0)	
Total, n (%)	1 (100.0)	1 (50.0)	
Patients with treatment changes, n (%)	4 (25)	18 (90.0)	$r = 0.14$
Patients with hospitalizations, n (%)	2 (12.5)	4 (20.0)	$r = 0.15$
Patients with ≥ 1 flare, n (%)	3 (18.8)	10 (50.0)	$r = 0.05$
No hospitalization, n (%)	2 (66.7)	8 (80.0)	
Hospitalization, n (%)	1 (33.3)	2 (20.0)	

RD, red density; IQR, interquartile range; NA, not applicable

Hence, the studied RD technology is currently being examined in a worldwide prospective multicenter trial (PROCEED-UC trial, clinicaltrials.gov identifier NCT 04408703) for its predictive value in patients with UC in clinical remission and its correlation with endoscopic and histologic scoring systems, as well as with biological markers (hemoglobin, platelet count, serum albumin, and fecal calprotectin). This ongoing trial includes patients with UC in clinical remission for at least 3 months without any recent (< 3 months) or planned therapy change. Follow-up is 52 weeks with a 13-week interval for clinical and biochemical check-up. At baseline and after 52 weeks, enrolled patients undergo flexible sigmoidoscopy for imaging in the rectum (5–10 cm) and sigmoid (30–35 cm) and biopsy (4 specimens) at corresponding imaging sites. MES and UCEIS scoring together with histological Geboes scoring are both performed for each endoscopy.

The current retrospective results show a promising trend for the predictive value of the RD score. However, interpretation of the current retrospective data has obvious limitations. First, this was a retrospective analysis of a prospective pilot trial, which was not statistically powered for prediction analysis, leading to potentially too little data for correct interpretation of these analyses. Second, all data from the patient cohort were collected at a single center tertiary IBD referral hospital, which may have led to a selection bias. Third, all patients were enrolled consecutively, resulting in a heterogenous group of different patients with different disease stages and treatments.

Hence, we emphasize that the further prospective PROCEED-UC trial with appropriate sample size and ideal target group is required to provide deeper insights about the predictive value of RD and the appropriate cut-off for prediction of disease course.

Conclusions

The RD score may be a considerable independent predictor of clinical remission in patients with UC because it tends to objectively indicate the disease course up to 5 years. A RD score > 65 demonstrated an increase in the composite endpoint. However due to limitations, this RD technology needs further investigation to draw definite and stable conclusions about its predictive capacity. Hence, results of the ongoing PROCEED-UC trial are to be awaited, but current trends look very promising.

Conflict of Interest

Funding

References

[1] Colombel JF, Rutgeerts P, Reinisch W et al. Early mucosal healing with infliximab is associated with improved long-term clinical outcomes in ulcerative colitis. Gastroenterology 2011; 141: 1194–1201

[2] Shah SC, Colombel J-F, Sands BE et al. Mucosal healing is associated with improved long-term outcomes of patients with ulcerative colitis: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2016; 14: 1245–1255

[3] Ardizzone S, Cassinotti A, Duca P et al. Mucosal healing predicts late outcomes after the first course of corticosteroids for newly diagnosed ulcerative colitis. Clinical Gastroenterol Hepatol 2011; 9: 483–489 doi:10.1016/j.cgh.2010.12.028

[4] Osada T, Ohkusa T, Yokoyama T et al. Comparison of several activity indices for the evaluation of endoscopic activity in UC: Inter- and intra-observer consistency. Inflamm Bowel Dis 2010; 16: 192–197

[5] Vuitton L, Peyrin-Biroulet L, Colombel JF et al. Defining endoscopic response and remission in ulcerative colitis clinical trials: an international consensus. Aliment Pharmacol Ther 2017; 45: 801–813 doi:10.1111/apt.13948

[6] 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: 13–19

[7] 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: 20–25

[8] 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: 1684–1692

[9] Bossuyt P, Nakase H, Vermeire S et al. Automatic, computer-aided determination of endoscopic and histological inflammation in patients with mild to moderate ulcerative colitis based on red density. Gut 2020; 69: 1778–1786

[10] Takenaka K, Ohtsuka K, Fujii T et al. Deep neural network accurately predicts prognosis of ulcerative colitis using endoscopic images. Gastroenterology 2021; 160: 2175–2177 doi:10.1053/j.gastro.2021.01.210