Digestive and Liver Disease 48 (2016) 703–708

Contents lists available at ScienceDirect

Digestive and Liver Disease

journal homepage: www.elsevier.com/locate/dld

Review Article

Endoscopic scores for inflammatory bowel disease in the era of 'mucosal healing': Old problem, new perspectives

Cristiano Pagnini[a,*], Francesca Menasci[a], Federico Desideri[a], Vito D. Corleto[b], Gianfranco Delle Fave[a], Emilio Di Giulio[b]

[a] "Sapienza" University of Rome, Faculty of Medicine and Psychology, Digestive and Liver Disease Unit, S. Andrea Hospital, Rome, Italy
[b] Digestive Endoscopy Unit, S. Andrea Hospital, Rome, Italy

ARTICLE INFO

Article history:
Received 28 January 2016
Accepted 7 March 2016
Available online 14 March 2016

Keywords:
Crohn's disease
Endoscopic score
Inflammatory bowel disease
Mucosal healing
Ulcerative colitis

ABSTRACT

The importance of the endoscopic evaluation in inflammatory bowel disease (IBD) management has been recognized for many years. However, the modalities for reporting endoscopic activity represent an ongoing challenge. To address this, several endoscopic scores have been proposed. Very few have been properly validated, and the use of such tools remains sub-optimal and is mainly restricted to clinical trials. In recent years, a growing emphasis of the concept of 'mucosal healing' as a prognostic marker and therapeutic goal has increased the need for a more accurate definition of endoscopic activity in both ulcerative colitis (UC) and Crohn's Disease (CD). In the present review, the evolution of the challenges related to endoscopic scores in IBD has been analyzed, with particular attention paid to the renewed relevance of endoscopic activity in recent years. Currently, despite the growing relevance of endoscopic activity, evaluating this activity in IBD is still a challenge. The implementation of efficacious endoscopic scores and a better definition of the absence of activity (mucosal healing) are needed.

1. Introduction

Inflammatory bowel disease (IBD) is a pathologic condition that causes chronic inflammation and may involve several tracts of the intestine. The relevance of the endoscopic evaluation in the management of IBD for diagnosis and follow-up is indisputable. This is particularly true for ulcerative colitis (UC), which is characterized by a mucosal inflammation that does not cross the lamina propria and that involves only the colon to varying degrees [1]. By contrast, in Crohn's Disease (CD) the inflammation is typically transmural and may involve several intestinal tracts, including areas not easily accessible by an endoscope [2].

Several scores for the evaluation of endoscopic activity in IBD have been proposed. Although the first attempts to establish a numeric score to evaluate the endoscopic activity of UC date back to the 1950s, the challenge of correctly interpreting and codifying endoscopic features has not been solved, and the issue has become more important over the last decade. The wider utilization of potent therapeutic options for IBD patients, such as biologic drugs, has led to a consistent conceptual modification of the clinical management, with a subsequent need for the identification of new end-points for the evaluation of therapeutic efficacy. This process has helped move endoscopic activity evaluation toward the more stringent concept of 'mucosal healing' (MH). In fact, a growing body of evidence suggests that the complete resolution of evident signs of inflammation during an endoscopic examination is associated with a better long-term outcome in terms of shorter hospital stays, an increase in sustained remission, and a decrease in the need for surgery [3]. The evidence for and limitations of the MH concept are still matter of debate in the literature [4] and are beyond the scope of the present paper. However, MH's growing relevance in the management of IBD has definitely influenced the challenges of endoscopic evaluation of disease activity.

Although different diseases have different issues, on a fundamental level an endoscopic score is needed to codify the different grades of intestinal disease activity observed during an endoscopic examination in a quantitative, simple, reproducible, and acceptable way. An early issue in the development of a scoring system has been the identification of high-risk lesions that are easily recognizable during an endoscopy. In addition, the increased complexity of the evaluation has led to growing disagreement between different operators (inter-observer variability) and even in repeated evaluations by the same operator (intra-observer variability). It is crucial to develop a score that will reduce such variation and standardize the interpretation of results in both trials and clinical practice

* Corresponding author at: "Sapienza" University of Rome, S. Andrea Hospital, Via di Grottarossa 1035, 00189 Rome, Italy. Tel.: +39 0633779; fax: +39 06 33776601.
E-mail address: cristiano.pagnini@uniroma1.it (C. Pagnini).

http://dx.doi.org/10.1016/j.dld.2016.03.006
1590-8658/

[5]. The gradual shift from a symptom-based to a mucosal-based approach for IBD has led to new interest in the issue of endoscopic scores and reconsideration of the finality of the score itself. There has been a transition from the target to the gradation of endoscopic severity and therefore to the discrimination of patients based on disease severity and identification of the subset of patients meeting the MH criteria, despite MH's variable definition. Such attempts cannot be finalized because no clear definition has been stated for MH in either UC or CD. For the latter disease, clinical trials generally refer to the absence of ulcer, while for UC, most trials refer to a Mayo endoscopic score of 0-1 [6]. The lack of a validated cut-off for endoscopic scores in CD and the limitation of the current Mayo definition of 0-1 in UC further confirm that MH is yet to be clearly defined. In UC in particular, the current belief that endoscopic score of 0 and 1 have the same prognostic relevance has recently been challenged [7,8], and the finding of a relatively higher number of patients with endoscopic healing, as previously defined, than with clinical remission in some trials, such as ULTRA and PURSUIT [9,10], has raised additional questions about the definition of MH. Currently, primary end-points in clinical trials are still represented by clinical parameters and composite indexes, including symptoms, signs and lab tests. The scarce applicability of such parameters in clinical practice is decisively pushing toward the identification and better clarification of appropriate new end-points, and among these, endoscopic disease activity evaluation is one objective parameter that should be included [11].

The relevance of a correct evaluation of disease activity and of agreement among operators has been strongly reinforced by a recent observation that different evaluations of endoscopic reports, such as the evaluation by a 'central reader', may profoundly alter the report of a trial. In a randomized, double-blind, placebo-controlled, multicenter study investigating the safety and efficacy of a delayed-released mesalamine formulation in UC patients, the primary end-point of clinical remission at week 6 was not achieved in the original intention-to-treat analysis (30 vs. 20.6% in the treated vs. placebo group, $p = 0.069$). Instead, it was observed after a central reading revision by seven expert readers who independently analyzed sigmoidoscopy recordings (29 vs. 13.8% in the treated vs. placebo group, $p = 0.011$) and excluded 87 of 281 patients (31%) from the original population because they did not meet the inclusion criteria [12].

2. Ulcerative colitis

The evolution of the concept of endoscopic score is particularly evident in UC because the role of the endoscopic activity in disease management has long been recognized. Numerous scores have been used for the assessment of endoscopic activity in UC, and a recent systematic review identified 31 scoring systems proposed in the literature [13]. Among these, very few were constantly utilized in clinical trials. Early in 1937, Bargen et al. described the importance of directly evaluating the mucosa of patients with ulcerative colitis using a rigid proctoscope in conjunction with a magnifying attachment [14], but the first endoscopic score to evaluate disease activity in UC patients in a clinical trial was proposed by Truelove and Witts in 1955 [15]. In 1964, Baron et al. developed a score to evaluate the appearance of rectosigmoid mucosa using a rigid proctoscope, in which disease activity was expressed by severity of bleeding and friability without considering ulcerations [16]. The Baron score was modified by Feagan et al., in 2005, with the removal of the qualitative assessment of different levels of bleeding, and the inclusion of ulceration in the evaluation [17]. Baron score has represented for a long time the most frequently used endoscopic score in clinical trials. Its modified version represents a simple and efficacious tool to evaluate endoscopic activity, but the lack of an appropriate validation and the unclear definition of MH may limit the utilization of the score in the future.

In 1987, Schroeder et al. described an instrument to measure disease activity called the Mayo score, which included both endoscopic and clinical items [18]. The part of this composite score that evaluate endoscopic activity, namely the Mayo endoscopic subscore (MES), is the scoring system most widely used in clinical trials to describe endoscopic activity as easily reproducible. It is characterized by 4 components and scores range from 0 to 3, with 0 indicating inactive disease with normal mucosa; 1 indicating mild disease (erythema and mild friability); 2 indicating moderate disease (marked erythema, friability, absent vascular pattern and erosions); and 3 indicating severe disease (spontaneous bleeding and diffuse ulceration). MH was defined as a MES of 0 or 1, which predicts a better outcome than a MES of 2 or 3 with an associated reduction in colectomy. MES has been used in multiple clinical trials [13]. Its main limitations are the fact that the definition of MH, although commonly used, has not been validated, and as mentioned above, recent data suggest a difference prognostic relevance between the scores of 0 and 1 [7]. Moreover, the score does not discriminate between superficial and deep ulcerations, and grades 1 and 2 may present significant overlap. Osada et al. found adequate inter- and intra-observer agreement from the experts, but agreement was markedly lower when trainees were involved in the scoring [19]. More recently, the suboptimal agreement of the score was confirmed, and surprisingly, the agreement coefficient was higher among non-expert than expert gastroenterologists (kappa = 0.71 vs. 0.53, respectively) [20].

The need for a better definition of endoscopic activity and the lack of a fully validated tool have recently led to the proposal of new indexes in an attempt to increased inter-observer agreement. In 2012, Travis et al. proposed the Ulcerative Colitis Endoscopic Index of Severity (UCEIS) [21], and, in a subsequent study, the authors evaluated the reliability of the UCEIS and validated it with an independent cohort of investigators [22]. The UCEIS score represents a validated index of endoscopic severity of UC, and it is considered a useful tool in clinical practice for reducing variations between different observers. However, it has several limitations. As for other indexes, the definition of MH remains unclear. In addition, the thresholds for mild to moderate and severe disease have not been completely validated. Only the most severely affected part of the mucosa is scored, meaning that disease extension is not evaluated. Most importantly, even though the score offers a detailed description of mucosal inflammation, a real advantage in the inter-observer agreement over simpler scores, such as Baron and MES, has yet to be demonstrated [12].

In an attempt to include disease extension and to increase inter-observer agreement, the Ulcerative Colitis Colonoscopic Index of Severity (UCCIS) was recently proposed [23]. The score shows significant correlation with the clinical indexes, laboratory measurement of active disease, and good correlation with patient-defined remission. Therefore, the UCCIS may represent a useful tool based on the evaluation of each colonic segment that provides reproducible results in endoscopic scoring of patients with UC. However, such a score presents some limitations. First, the original study for the development of the score was a single-center study of only 50 patients, and larger studies are needed to validate the usefulness of this index. Second, calculating the score requires a total colonoscopy, which is a more expensive and invasive procedure that is less well tolerated by the patient compared to sigmoidoscopy. Moreover, UCCIS is pretty more complex comparing with other scores and therefore its widespread utilization in every endoscopic center may not be feasible.

Very recently, to conjugate the simplicity of the MES with the need to evaluate disease extension, a Modified Mayo Endoscopic Score (MMES) was proposed [24]. Lobaton et al. calculated this new