Cochrane
Library

Cochrane Database of Systematic Reviews

Endoscopic scoring indices for evaluation of disease activity in Crohn's disease (Review)

Khanna R, Nelson SA, Feagan BG, D'Haens G, Sandborn WJ, Zou GY, MacDonald JK, Parker CE, Jairath V, Levesque BG

Khanna R, Nelson SA, Feagan BG, D'Haens G, Sandborn WJ, Zou GY, MacDonald JK, Parker CE, Jairath V, Levesque BG.
Endoscopic scoring indices for evaluation of disease activity in Crohn's disease.
Cochrane Database of Systematic Reviews 2016, Issue 8. Art. No.: CD010642.
DOI: [10.1002/14651858.CD010642.pub2](https://doi.org/10.1002/14651858.CD010642.pub2).

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	3
OBJECTIVES	3
METHODS	3
RESULTS	5
Figure 1.	6
DISCUSSION	10
AUTHORS' CONCLUSIONS	11
ACKNOWLEDGEMENTS	11
REFERENCES	12
CHARACTERISTICS OF STUDIES	15
ADDITIONAL TABLES	33
APPENDICES	51
CONTRIBUTIONS OF AUTHORS	52
DECLARATIONS OF INTEREST	52
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	54
INDEX TERMS	54

[Methodology Review]

Endoscopic scoring indices for evaluation of disease activity in Crohn's disease

Reena Khanna^{1,2}, Sigrid A Nelson³, Brian G Feagan^{1,2,4,5}, Geert D'Haens^{6,7}, William J Sandborn^{3,8}, GY Zou^{2,5}, John K MacDonald^{1,4}, Claire E Parker², Vipul Jairath^{1,2,5}, Barrett G Levesque³

¹Department of Medicine, University of Western Ontario, London, Canada. ²Robarts Clinical Trials, London, Canada. ³Robarts Clinical Trials, San Diego, CA, USA. ⁴Cochrane IBD Group, Robarts Clinical Trials, London, Canada. ⁵Department of Epidemiology and Biostatistics, University of Western Ontario, London, Canada. ⁶Academic Medical Center, Amsterdam, Netherlands. ⁷Robarts Clinical Trials, Amsterdam, Netherlands. ⁸Division of Gastroenterology, University of California San Diego, La Jolla, CA, USA

Contact address: Brian G Feagan, Cochrane IBD Group, Robarts Clinical Trials, 100 Dundas Street, Suite 200, London, ON, N6A 5B6, Canada. brian.feagan@robartsinc.com.

Editorial group: Cochrane IBD Group.

Publication status and date: New, published in Issue 8, 2016.

Citation: Khanna R, Nelson SA, Feagan BG, D'Haens G, Sandborn WJ, Zou GY, MacDonald JK, Parker CE, Jairath V, Levesque BG. Endoscopic scoring indices for evaluation of disease activity in Crohn's disease. *Cochrane Database of Systematic Reviews* 2016, Issue 8. Art. No.: CD010642. DOI: [10.1002/14651858.CD010642.pub2](https://doi.org/10.1002/14651858.CD010642.pub2).

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Endoscopic assessment of mucosal disease activity is widely used to determine eligibility and response to therapy in clinical trials of treatment for Crohn's disease. However, the operating properties of the currently available endoscopic indices remain unclear.

Objectives

A systematic review was undertaken to evaluate the development and operating characteristics of the Crohn's Disease Endoscopic Index of Severity (CDEIS) and Simple Endoscopic Scale for Crohn's Disease (SES-CD).

Search methods

Electronic searches of the MEDLINE (1966 to December 2015), EMBASE (1980 to December 2015), and Cochrane CENTRAL Register of Controlled Trials (Issue 12, 2015) databases were supplemented by manual reviews of reference listings and conference proceedings (Digestive Disease Week, United European Gastroenterology Week, European Crohn's and Colitis Organization).

Selection criteria

Any study design (e.g. randomized controlled trials, cohort studies, case series) that evaluated either or both the CDEIS or SES-CD in patients with Crohn's disease was considered for inclusion. Eligible participants were adult patients (≥ 16 years), diagnosed with Crohn's disease using conventional clinical, radiographic, and endoscopic criteria.

Data collection and analysis

Two authors (RK, JKM) independently reviewed the titles and abstracts of the studies identified from the literature search. The full texts of potentially relevant citations were reviewed for inclusion and the study investigators were contacted to clarify any unclear data. Any disagreements were resolved by discussion and consensus with a third author. A standardized form was used to assess eligibility of trials for inclusion in the study and for data extraction.

Two authors independently extracted and recorded data (RK, SAN). The number of patients enrolled; number of patients per treatment arm; patient characteristics including age and gender distribution; endoscopic index; and outcomes such as intra-rater reliability, inter-rater reliability responsiveness, validity, feasibility, construct validity, and criterion validity were recorded for each trial.

Main results

Forty-three reports of 30 studies fulfilled the inclusion criteria.

For the SES-CD, inter-rater reliability was assessed in four studies. In the development study for the SES-CD ([Daperno 2004](#)), the overall ICC (0.9815, 95% CI 0.9705 to 0.9884) and the kappa for the regions is high; however the paired raters were in the same room which introduces the potential for bias.

For the CDEIS, inter-rater reliability was assessed in six studies. [Daperno 2014](#) reported that the ICC for the CDEIS was 0.985 (95% CI 0.939-1.000) for average measures of video score and was 0.835 (95% CI 0.540-0.995) for single measures of video score.

With respect to validity, correlation between the CDEIS and clinical measures, including C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR), was also reported. The estimates of correlation with CRP were $r = 0.521$ ([Sipponen 2010b](#)), $r = 0.553$ ([Sipponen 2008a](#)) and $r = 0.608$ ([Sipponen 2008c](#)). For the SES-CD, the corresponding values for correlation with CRP ranged from $r = 0.46$ ([Jones 2008](#)) to $r = 0.68$ ([Green 2011](#)).

Responsiveness data for the CDEIS were available in nine studies. Seven studies demonstrated statistically significant decreases in the CDEIS score after administration of a treatment of known efficacy. Minimal responsiveness data were available for the SES-CD. [Sipponen 2010a](#) and [Sipponen 2010b](#) demonstrated statistically significant changes in the SES-CD score after subjects were administered a treatment of known efficacy.

No studies were identified that explicitly evaluated the feasibility for either the SES-CD or the CDEIS. The SES-CD requires fewer calculations and may therefore be easier to use than the CDEIS.

Authors' conclusions

Although they are used in clinical trials, the CDEIS and SES-CD remain incompletely validated. Future research is required to determine the operating properties and to define the optimal index.

PLAIN LANGUAGE SUMMARY

Endoscopic scoring indices for evaluation of disease activity in Crohn's disease

What is Crohn's disease?

Crohn's disease is a long-term (chronic) inflammatory bowel disease characterized by pain (abdominal cramping), tiredness, diarrhea and weight loss. When people with Crohn's disease are experiencing symptoms the disease is said to be "active" and when symptoms stop this is called "remission".

What is an endoscopic scoring index?

An endoscopic scoring index measures disease activity based on what a doctor sees during endoscopy. An endoscopy is a non-surgical procedure used to view the digestive tract using a camera. The doctor who performs the endoscopy can rate disease activity using the index, or this can be done by another off-site doctor if a video of the procedure was recorded.

The most commonly used endoscopic scoring indices are the Crohn's Disease Endoscopic Index of Severity (CDEIS) and the Simple Endoscopic Score for Crohn's Disease (SES-CD).

What did the researchers investigate?

It is important that endoscopic scoring indices are valid, meaning that they accurately measure what they are supposed to measure. The researchers investigated whether studies have assessed the validity of the CDEIS and/or the SES-CD.

What did the researchers find?

The researchers found that neither the CDEIS nor the SES-CD has been fully validated.

BACKGROUND

Description of the problem or issue

Crohn's disease (CD) is a chronic inflammatory disease of unknown etiology that can affect any part of the gastrointestinal tract. Clinically, the disease presents with diarrhea, abdominal pain, fatigue, and weight loss and is characterized a relapsing-remitting course. Endoscopic findings include mucosal edema, erythema granularity, friability, and ulcers. Disease activity may be classified as mild, moderate, severe, or fulminant based on combined clinical and endoscopic assessments. Although CD may involve any part of the gastrointestinal tract, approximately 20% of patients have isolated colonic involvement, 30% have disease limited to the small bowel, and 50% of patients have involvement of both the ileum and the colon.

Several endoscopic and clinical indices have been developed to evaluate CD activity in clinical trials. It is now apparent that a poor correlation exists between symptoms, as measured by an instrument such as the Crohn's Disease Activity Index (CDAI), and endoscopic measures (Cellier 1994; Regueiro 2011). Recently, increased importance has been placed on the use of endoscopic instruments as outcome measures in clinical research, since they are potentially more objective than symptom-based indices (FDA 2013). Therefore, the operating properties of these endoscopic indices must be clearly defined. In particular, an endoscopic index must be valid (it must measure the outcome that it is intended to assess); responsive (it must detect a meaningful change in health status); and reliable (consistent results are obtained in patients whose clinical status has remained stable) (Kirshner 1985). Furthermore, an ideal instrument should be highly feasible to use in the setting of a clinical trial. Currently, the Crohn's Disease Endoscopic Index of Severity (CDEIS) (Mary 1989) and the Simple Endoscopic Score for Crohn's Disease (SES-CD) (Daperno 2004) are the most commonly used indices to assess eligibility and response to therapy in clinical trials.

The CDEIS assesses five segments of the intestine: rectum, sigmoid and left colon, transverse colon, right colon, and ileum (Mary 1989). For each colonic segment, the presence of deep and superficial ulceration, the percentage of ulcerated surface and the percentage of surface involved by CD are measured on a 10 centimetre visual analogue scale. Additionally, the presence of ulcerated and non-ulcerated stenoses are assessed. These items are weighted and summed to yield a total score that ranges from 0 to 44. Higher scores indicate more severe disease.

The SES-CD evaluates four endoscopic items (ulcer size, proportion of the surface area that is ulcerated, proportion of the surface area affected, and stenosis). Each item is scored by segment on a scale from zero to three (Daperno 2004). The maximum SES-CD score is 56, with higher scores indicating more severe disease.

Why it is important to do this review

There are few data available on the operating properties of these two indices despite their widespread use as outcomes in clinical trials. This review will evaluate the relative merits of the two indices and identify areas where further validation of the CDEIS and SES-CD is required.

OBJECTIVES

The primary objective is to systematically review the current literature describing the development and operating characteristics of endoscopic disease activity indices used to evaluate the activity of Crohn's disease.

METHODS

Criteria for considering studies for this review

Types of studies

Any study design (e.g. randomized controlled trials, cohort studies, case series) that evaluates either the SES-CD or CDEIS or both in patients with Crohn's disease was considered for inclusion. Study subjects included adult patients (≥ 16 years), diagnosed with Crohn's disease using conventional clinical, radiographic, and endoscopic criteria.

Types of data

Endoscopic scoring data obtained from eligible studies were considered for inclusion.

Types of methods

The methods used to construct and validate the endoscopic indices (e.g. reliability, validity, responsiveness and feasibility) were examined in detail and described for each eligible study. We also reported the number of endoscopists who scored the indices in each study and whether they were blinded to other raters' scores.

Types of outcome measures

Reliability: For each study reliability was assessed by measuring intra-rater reliability, inter-rater reliability or internal consistency. Intra- and inter-rater reliability were assessed by using the interclass correlation coefficient (ICC) for repeat assessments made by the same rater, and for assessments made by different raters.

Validity: Validity was broadly defined as evidence that variations in Crohn's disease severity causally produce variations in the endoscopic indices' measurement outcomes. Studies were assessed to determine whether content validity, criterion validity, or construct validity for endoscopic index scores in specific clinical situations were reported.

A Crohn's disease endoscopic index demonstrates content validity if the components of the index are sufficient to measure disease activity. Content validation was generally based on qualitative assessment. For example, evidence of content validity includes expert panel opinion on face validity, or a systematic review of the literature supporting the development of an endoscopic index.

Criterion validity was demonstrated if the index scores are an adequate reflection of true Crohn's disease activity as assessed by a gold standard. The lack of a single gold standard for Crohn's disease activity is a limitation of these assessments. Studies of predictive criterion validity will be reported which compare whether the score predicts true Crohn's disease activity as measured by objective measures of inflammation including C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), fecal lactoferrin, or fecal calprotectin. Studies of criterion validity which compare endoscopic scores with future sequelae, such as surgery or disability will also be reported. Statistical parameters for

agreement between the “EI and gold standards” were assessed (i.e., sensitivity or specificity, receiver operating characteristic (ROC) curve, area under the receiver operating characteristic curve (AUC), mean difference, weighed kappa, Spearman's r squared, and the ICC).

Studies that assess construct validity of the endoscopic indices, which acknowledge the limitation of the lack of a gold standard for disease activity and assess whether the endoscopic indices are consistent with other hypotheses of true disease activity were included. Studies of construct validity of the endoscopic index were reported if they measured the correlation between the scores obtained from endoscopic and clinical indices, including the physician's global assessment of disease activity. The minimal clinically important difference (MCID) and appropriate cut-off values to determine active and inactive disease states, clinical response, and clinical remission were examined.

Responsiveness: Responsiveness, the ability of the index to detect change, was determined by assessing changes in the endoscopic index scores following a treatment of known efficacy. Responsiveness was quantified using indicators of effect size ([Zou 2005](#)), or with the use of ROC curves to describe how well score changes can distinguish improved patients from those without response ([Deyo 1991](#)).

Feasibility: Feasibility was assessed as the raters' evaluation of the ease of administration and time required for scoring.

Search methods for identification of studies

Electronic searches

We searched the following databases from inception to 1 December 2015:

1. MEDLINE;
2. EMBASE;
3. The Cochrane Library (CENTRAL); and
4. The Cochrane IBD Group Specialized Register.

The search strategies for each database are documented in [Appendix 1](#).

Searching other resources

We also searched major gastroenterology meetings (2000 to present) including:

1. Digestive Disease Week (DDW);
2. United European Gastroenterology Week (UEGW); and
3. European Crohn's and Colitis Organization (ECCO).

Reference lists from retrieved articles were scanned to identify additional citations that may have been overlooked by the database search.

Data collection and analysis

Selection of studies

Two authors (RK and JKM) independently reviewed the titles and abstracts of the studies identified from the literature search. The full text of potentially relevant citations were reviewed for inclusion and the study investigators were contacted to clarify any unclear data, as required. Any disagreements were resolved by discussion and consensus with a third author (BGF).

A standardized form was used to assess eligibility of trials for inclusion in the study. Each item on the form was scored as yes, no or unclear. The following items were assessed:

- a) Diagnosis of Crohn's disease; and
- b) Assessment of CDEIS or SES-CD.

Data extraction and management

A standardized form was used to extract information from selected studies. Two authors (RK, SAN) independently extracted and recorded data. The following data were recorded from each eligible study:

- a) Number of patients enrolled, number of patients per treatment arm;
- b) Patient characteristics: age and gender distribution;
- c) Endoscopic index: CDEIS or SES-CD;
- d) Thresholds in the CDEIS or SES-CD associated with clinical response or remission; and
- e) Outcomes: intra-rater reliability; inter-rater reliability; responsiveness; validity; feasibility; construct validity; criterion validity, and the MCID.

Assessment of risk of bias in included studies

We used the following criteria to appraise the risk of bias of included studies:

- Blinding to clinical information
- Independent observation by endoscopists

Blinding to clinical information such as symptoms, physical examination, or laboratory information is important for objective assessment of endoscopic data ([Feagan 2013](#)). Furthermore, independent observation is essential to ensure confidence in the inter-rater reliability coefficients.

We also assessed the methodological quality of the included studies using the COSMIN (Consensus-based Standards for the selection of health Measurement INstruments) checklist. The checklist consists of 10 properties: internal consistency, reliability, measurement error, content validity, structural validity (factor analysis), hypothesis testing, cross-cultural validity, criterion validity, responsiveness to change and interpretability. Each property is rated on a four-point scale (1 = poor, 2 = fair, 3 = good, or 4 = excellent). The overall score for the assessment of an individual measurement property is obtained by taking the lowest score for any of the items in the property box (i.e. if any item in the property

box is scored as "poor" then the overall score for that property is "poor").

Measures of the effect of the methods

Descriptive statistics were used to report the validation of outcome data. Frequencies and percentages were shown for categorical variables.

Dealing with missing data

Where possible, authors were contacted to provide any missing information.

RESULTS

Description of studies

Results of the search

The literature search conducted on 1 December 2015 identified 4334 records. After duplicates were removed, a total of 3039 records were screened for inclusion. Of these, 55 reports were selected for full text review. Ten reports of seven studies were excluded (see [Characteristics of excluded studies](#)), leaving 45 reports of 31 studies that met the pre-defined inclusion criteria (see [Figure 1](#)).

Figure 1. Study flow diagram.

Included studies

Thirty studies (43 reports) were completed at the time of screening and reported validation results (Af Borjorkesten 2012; Bamba 2012; Colombel 2003; D'Haens 1999; Daperno 2004; Daperno 2014; Denis 2007; Ferrante 2010; Ferrante 2013; Green 2011; Hebutterne 2013; Jalocha 2009; Jones 2008; Khanna 2016; Lobaton Ortega 2013; Mary 1989; Molander 2013; Reinisch 2010; Rimola 2011; Rutgeerts 2002; Rutgeerts 2006; Rutgeerts 2010; Rutgeerts 2016; Schoepfer 2010; Sipponen 2008a; Sipponen 2008b; Sipponen 2008c; Sipponen 2010a; Sipponen 2010b; Vieira 2009). One study (two reports) was ongoing and reported preliminary results but no validation results (Buisson 2015).

Of the 30 included studies, 10 evaluated the SES-CD (Af Borjorkesten 2012; Bamba 2012; Green 2011; Jones 2008; Molander 2013; Reinisch 2010; Rutgeerts 2010; Schoepfer 2010; Sipponen 2008b; Sipponen 2010a), 12 evaluated the CDEIS (Colombel 2003; D'Haens 1999; Denis 2007; Hebutterne 2013; Jalocha 2009; Mary 1989; Rimola 2011; Rutgeerts 2002; Rutgeerts 2006; Sipponen 2008a; Sipponen 2008c; Vieira 2009), and 8 evaluated both the SES-CD and CDEIS (Daperno 2004; Daperno 2014; Ferrante 2010; Ferrante 2013; Khanna 2016; Lobaton Ortega 2013; Rutgeerts 2016; Sipponen 2010b). See Table 1.

Various thresholds were used to define response and remission using the SES-CD or CDEIS in the included studies. A summary of the thresholds used by each study is provided in Table 2.

Setting

Three studies were based on prospective assessment during an endoscopy procedure with two raters viewing at the same time (Af Borjorkesten 2012; Daperno 2004; Mary 1989). Similarly, 16 studies were based on prospective assessment during an endoscopy procedure, but with a single rater (Colombel 2003; D'Haens 1999; Denis 2007; Jalocha 2009; Jones 2008; Lobaton Ortega 2013; Molander 2013; Rimola 2011; Rutgeerts 2002; Rutgeerts 2006; Schoepfer 2010; Sipponen 2008a; Sipponen 2008b; Sipponen 2008c; Sipponen 2010a; Sipponen 2010b). Seven studies were based on retrospective review of endoscopy videos (Daperno 2014; Ferrante 2010; Ferrante 2013; Hebutterne 2013; Khanna 2016; Rutgeerts 2016; Vieira 2009). For Bamba 2012 and Green 2011, retrospective endoscopy rating was performed based on material in medical charts, however it was not clear how this was performed. Reinisch 2010 and Rutgeerts 2010 did not specify how the assessments were performed.

Mary 1989 described the development of the CDEIS in a multi-phase study. During the development phase, colonoscopies from 75 patients were assessed by two endoscopists in the same room. Data on mucosal lesions, the percentage of segmental surfaces with disease involvement, and the percentage of segmental surfaces with ulcerations were evaluated in five endoscopic segments (rectum, sigmoid and left colon, transverse colon, right colon, and ileum). Multiple linear regression identified independent variables correlated with the Global Evaluation of Lesion Severity (GELS) to generate an overall score. In the next phase of the study, agreement was assessed based on the intra-class correlation coefficient (ICC) and the correlation between the CDEIS with the GELS was used to determine criterion validity. In the final phase, 54 patients with active Crohn's disease underwent colonoscopies at baseline

and three to five weeks after oral prednisone (1mg/kg/day) to determine the responsiveness of the CDEIS.

The SES-CD was originally constructed in a four part study (Daperno 2004). In part one, the reproducibility of the several items were examined. Two endoscopists scored 71 colonoscopies to assess ulcer size, the proportion of the surface with ulcers, the proportion of the mucosal surface affected with disease, and stenosis in five endoscopic segments (rectum, sigmoid and left colon, transverse colon, right colon, and ileum). Part two of the study developed the SES-CD using multiple linear regressions of items with strong reproducibility. As a result of this process, the SES-CD was defined as the sum of four variables over five colonic segments. The validation phase comprised part three of the study, in which 121 additional patients were included. The statistical correlation between the SES-CD and the CDEIS was calculated. The fourth part of this study evaluated the reliability of the SES-CD and CDEIS, including correlation with clinical parameters such as the CDAI, the Inflammatory Bowel Disease Questionnaire (IBDQ), and biochemical markers (CRP and albumin).

Excluded studies

Seven studies (Bondjemah 2012; Mantzaris 2009; Modigliani 1987; Sandborn 2012a; Sandborn 2012b; Sandborn 2012c; Ziech 2012), were excluded with reasons (see Characteristics of excluded studies).

Risk of bias in included studies

In Daperno 2004 multiple raters were in the same room during endoscopy to perform mucosal assessments. The authors state that independence of observation was upheld, but it is plausible that non-verbal communication could influence scoring yielding high agreement (e.g. the endoscopist pausing at an ulcer or facial reactions could bias the observer).

Similarly, multiple raters were in the same room in Mary 1989. The authors state the endoscopist and the observer did not communicate during the colonoscopic examination, and they completed the endoscopic scoring sheets independently.

In many studies a single rater assessed endoscopic disease activity (D'Haens 1999; Denis 2007; Ferrante 2010; Jalocha 2009; Molander 2013; Reinisch 2010; Rimola 2011; Rutgeerts 2002; Rutgeerts 2016; Schoepfer 2010; Sipponen 2008a; Sipponen 2008b; Sipponen 2008c; Sipponen 2010a; Sipponen 2010b). As a result rater independence was not relevant, but differences in training and expertise could cause variation. Additionally, results for a single expert reviewer may not be generalizable to other assessors. Several studies did not describe the process for selection of readers or the number of readers who evaluated each endoscopy. It is possible that a single reader assessed the endoscopy, or that independence of observation was not considered in the trial design (Bamba 2012; Colombel 2003; Ferrante 2013; Green 2011; Jones 2008; Lobaton Ortega 2013; Reinisch 2010; Rutgeerts 2010).

In studies with multiple readers assessing each endoscopy that did not report inter-rater reliability there is a potential concern that readers were not consistent (Vieira 2009). The process of reader training and testing, to ensure consistency of assessments between readers, was unclear in these circumstances.

Blinding

Blinding to clinical information such as symptoms, physical examination or laboratory information is important for the objective assessment of endoscopic data (Feagan 2013). The presence or absence of blinding was not routinely reported.

In 10 studies, raters were blinded to clinical information (Daperno 2014; Ferrante 2010; Ferrante 2013; Green 2011; Hebuterne 2013; Khanna 2016; Rimola 2011; Rutgeerts 2006; Rutgeerts 2016; Vieira 2009). In three studies, raters appeared to be partially blinded to clinical information (Sipponen 2008b; Jones 2008; Af Borjorkesten 2012). In 12 studies, it was not clear if assessors were blinded to clinical information (Colombel 2003; Jalocha 2009; Lobaton Ortega 2013; Mary 1989; Reinisch 2010; Rutgeerts 2002; Rutgeerts 2010; Schoepfer 2010; Sipponen 2008a; Sipponen 2008c; Sipponen 2010a; Sipponen 2010b). Finally, in five studies, it appeared that raters were not blinded to clinical information (Bamba 2012; D'Haens 1999; Daperno 2004; Denis 2007; Molander 2013). In studies where the assessment was performed by a local endoscopist, it is likely that the rater would have access to clinical data.

Results from the COSMIN assessment of methodological quality are reported in the [Effects of methods](#) section.

Incomplete outcome data

Differences in validity comparisons observed for the both the SES-CD and CDEIS (such as variation observed between correlation to CRP and fecal calprotectin) may be due to the heterogenous populations studied in terms of severity and history of CD as well as length of follow-up. Thus, no meta-analysis or summary measures of correlation were reported in this systematic review due to the differences in study designs.

Reliability estimates may also be influenced by sample selection and exclusions. Studies that did not include the full spectrum of disease activity could have improved ICC estimates. For example, a sample with primarily extreme populations (complete remission or severe CD) are less subjective to score consistently. Therefore, this may be one reason for differences observed in inter-rater agreement.

Selective reporting

Our search included a review of the grey literature and hand-searching of abstracts, but there is potential that unpublished validation data were not included in this review. However, there is limited risk that a major validation study was missed.

Other potential sources of bias

No other potential sources of bias were identified.

Effect of methods

Reliability

One study assessed intra-rater reliability (agreement with self over multiple assessments of the same endoscopy video) for both the SES-CD and CDEIS (Khanna 2016). The overall ICCs were 0.91 (95% CI 0.89 to 0.95) and 0.89 (95% CI 0.86 to 0.93) for the SES-CD and CDEIS, respectively (see [Table 3](#)).

For the SES-CD, inter-rater reliability (agreement between raters) was assessed in four studies (Daperno 2004; Daperno 2014; Khanna 2016; Rutgeerts 2016; see [Table 3](#)). Daperno 2004 was the original validation paper for the SES-CD, and provided inter-rater reliability data using observer pairs. Both the ICC for the total score and kappa values for components of the SES-CD score were reported ([Table 3](#)). The overall ICC is 0.9815, 95% CI 0.9705 to 0.9884 and the kappa for the regions is high, however the paired raters were located in the same room. The publication states there was no communication among endoscopists and that scoring sheets were completed independently, however there exists a potential for bias as described in the [Assessment of risk of bias in included studies](#) section. Daperno 2014 examined eight endoscopic videos with 13 raters, and also found high ICC (> 0.9) ([Table 3](#)). Rutgeerts 2016 observed overall ICCs of 0.77 (95% CI 0.69 to 0.83) at baseline, 0.86 (0.81 to 0.90) at week 12, and 0.82 (95% CI 0.73 to 0.88) at week 52.

For the CDEIS, inter-rater reliability was assessed in six studies (Daperno 2004; Daperno 2014; Hebuterne 2013; Khanna 2016; Mary 1989; Rutgeerts 2016 see [Table 3](#)). Mary 1989 was the original validation paper for the CDEIS and provided inter-rater reliability data using observer pairs, where the observer viewed through a teaching tube. Daperno 2004 compared the SES-CD to the CDEIS and found the ICC for the CDEIS was 0.9090 (95% CI 0.8580-0.9423) and also reported kappa values for components of the CDEIS score, which tended to be lower than those reported for the SES-CD. Daperno 2014 reported for the CDEIS the ICC was 0.985 (95% CI 0.939-1.000) for average measures and was 0.835 (95% CI 0.540-0.995) for single measures. In Hebuterne 2013 blinded and independent reviews of endoscopic videos were assessed in duplicate. The ICCs were 0.60 (95% CI 0.39-0.75) at baseline, 0.74 (95% CI 0.59-0.84) at week 10, and 0.81 (95% CI 0.65-0.90) at week 54. Rutgeerts 2016 reported the overall ICCs to be 0.78 (95% CI 0.65 to 0.7) at baseline, 0.92 (95% CI 0.86 to 0.95) at week 12, and 0.86 (95% CI 0.73 to 0.93) at week 52.

Although higher numerical values were observed for the SES-CD compared to the CDEIS, both indices demonstrated good inter-rater reliability in these limited number of studies.

Validity

Content validity

No studies assessed the content validity of the SES-CD or CDEIS (see [Table 4](#))

Criterion validity

For both the SES-CD and CDEIS, correlation between endoscopic score and clinical measures, including CRP, ESR, fecal calprotectin, and fecal lactoferrin were reported (see [Table 5](#)). The correlations ranged from poor to excellent, with the highest correlations observed between CDEIS and fecal calprotectin ($r = 0.831$) and CDEIS and fecal lactoferrin ($r = 0.865$).

CRP

For the SES-CD, the estimates of correlation with CRP ranged from $r = 0.46$ (Jones 2008) to $r = 0.68$ (Green 2011). Six additional studies published correlations within this range (Af Borjorkesten 2012; Bamba 2012; Daperno 2004; Jones 2008; Schoepfer 2010; Sipponen 2010b). The correlation between change in CRP and change in

SES-CD was reported as $r = 0.28$ by [Ferrante 2013](#), and was not statistically significant in [Sipponen 2010b](#).

For the CDEIS, the estimates of correlation with CRP were $r = 0.521$ ([Sipponen 2010b](#)), $r = 0.553$ ([Sipponen 2008a](#)) and $r = 0.608$ ([Sipponen 2008c](#)). One study reported that a correlation between CDEIS and CRP did not exist ([Denis 2007](#)). The correlation between the change in CRP and change in CDEIS was reported as $r = 0.27$ by [Ferrante 2013](#), $r = 0.46$ by [D'Haens 1999](#) and was not statistically significant in [Sipponen 2010b](#).

In [Lobaton Ortega 2013](#) the correlation between endoscopic activity and CRP was reported as $r = 0.650$, however it is unclear whether endoscopic activity was defined using the SES-CD or CDEIS.

ESR

For the SES-CD, [Bamba 2012](#) reported that the estimate of correlation with ESR was $r = 0.399$. [Green 2011](#) also explored the correlation between SES-CD and ESR and found that the two did not correlate ($p = 0.6$), but did not report the r value.

No studies explicitly investigated the correlation between CDEIS and ESR. In [Lobaton Ortega 2013](#) the correlation between endoscopic activity and ESR was reported as $r = 0.397$, however it is unclear whether endoscopic activity was defined using the SES-CD or CDEIS.

Fecal calprotectin

For the SES-CD, the estimates of correlation with fecal calprotectin were $r = 0.45$ ([Jones 2008](#)), $r = 0.525$ ([Sipponen 2010b](#)), $r = 0.56$ ([Af Bjorjorkesten 2012](#)), and $r = 0.75$ ([Schoepfer 2010](#)). [Sipponen 2008b](#) reported a correlation of $r = 0.642$ for patients with ileocolonic or colonic disease, and $r = 0.317$ for patients with disease affecting the ileum or upper gastrointestinal tract. No r values were reported in [Sipponen 2010a](#), however patients with a SES-CD defined response had a mean decrease in fecal calprotectin (from 1282 mg/g to 73 mg/g, $p = 0.005$). Patients without a response, as defined by SES-CD, had a lower mean decrease in fecal calprotectin (from 1017 mg/g to 223 mg/g, $p = 0.594$).

For the CDEIS, estimates of correlation with fecal calprotectin were reported as $r = 0.729$ in [Sipponen 2008a](#) and $r = 0.831$ in [Sipponen 2008c](#). In [Denis 2007](#) no correlation was detected, but the authors failed to report an r value.

In [Lobaton Ortega 2013](#) the correlation between endoscopic activity and fecal calprotectin was $r = 0.807$, however it is unclear whether endoscopic activity was defined using the SES-CD or CDEIS.

Fecal lactoferrin

For the SES-CD, the estimates of correlation with fecal lactoferrin were $r = 0.672$ for patients with ileocolonic or colonic disease, and $r = 0.180$ for patients with disease affecting the ileum and or upper gastrointestinal tract ([Sipponen 2008b](#)). The later correlation was not statistically significant. No r value was reported in [Sipponen 2010a](#), however patients with an SES-CD defined response had a mean decrease in fecal calprotectin from 223 mg/g to 0.6 mg/g ($p = 0.005$). Patients without an SES-CD defined response had a mean

decrease in fecal calprotectin from 22.5 mg/g to 13.0 mg/g ($p = 0.515$).

For the CDEIS, estimates of correlation with fecal lactoferrin were $r = 0.773$ in [Sipponen 2008a](#) and $r = 0.865$ in [Sipponen 2008c](#).

Construct validity

With regard to construct validity, the correlations ranged from poor to excellent. The highest correlations observed between the SES-CD and CDAI ($r = 0.92$), CDEIS and CDAI ($r = 0.92$), CDEIS and the Global Evaluation of Lesion Severity (GELS) ($r = 0.83$) and the CDEIS and SES-CD ($r = 0.968$).

For the SES-CD, estimates of correlation with the CDAI ranged from $r = 0.15$ ([Jones 2008](#)) to $r = 0.92$ ([Khanna 2016](#)). Five additional studies published correlations within this range ($r = 0.4$ [Af Bjorjorkesten 2012](#); $r = 0.39$ [Daperno 2004](#); $r = 0.38$ [Schoepfer 2010](#); $r = 0.346$ [Sipponen 2008b](#); and $r = 0.473$ [Sipponen 2010b](#)). In [Ferrante 2010](#) the estimate of correlation between mean change in SES-CD and CDEIS was $r = 0.27$. Two studies reported estimates of correlation between the SES-CD and Harvey-Bradshaw Index (HBI) as $r = 0.32$ ([Af Bjorjorkesten 2012](#)) and $r = 0.70$ ([Green 2011](#)). In [Bamba 2012](#) the estimate of correlation between the SES-CD and Modified Rutgeerts Score was $r = 0.748$. See [Table 6](#).

For the CDEIS, estimates of correlation with CDAI were reported as $r = 0.92$ in [Khanna 2016](#), $r = 0.442$ in [Lobaton Ortega 2013](#), $r = 0.381$ in [Sipponen 2008a](#), $r = 0.592$ in [Sipponen 2008c](#), $r = 0.446$ in [Sipponen 2010b](#). An estimate of correlation between mean change in CDEIS and CDAI was reported as $r = 0.22$ in [Ferrante 2013](#), $r = 0.307$ at week 10 and $r = 0.498$ at week 54 in [Rutgeerts 2002](#), and $r = 0.33$ at week 10, and $r = 0.37$ at week 54 in [Rutgeerts 2006](#). [Colombel 2003](#) and [Denis 2007](#) provided no r values, however the former found the CDEIS and CDAI to be positively correlated while the latter found no correlation between the two indices. [Khanna 2016](#) and [Mary 1989](#) provided estimates of correlation between the CDEIS and the Global Evaluation of Lesion Severity (GELS) score of $r = 0.74$ and $r = 0.83$, respectively. [Table 6](#).

Five studies explored the relationship between the CDEIS and SES-CD. Estimates of correlation were reported as $r = 0.357$ in [Daperno 2004](#), $r = 0.810$ in [Ferrante 2010](#), $r = 0.968$ in [Lobaton Ortega 2013](#) and $r = 0.938$ in [Sipponen 2010b](#). An estimate of correlation between mean change in CDEIS and SES-CD was reported as $r = 0.89$ in [Ferrante 2013](#). See [Table 6](#).

Responsiveness

Minimal responsiveness data are available for the SES-CD. [Sipponen 2010a](#) and [Sipponen 2010b](#) demonstrated statistically significant changes in the SES-CD score after subjects were administered a treatment of known efficacy. [Ferrante 2013](#) measured change after treatment, however statistics were only reported for a composite definition of corticosteroid-free endoscopic response, not for individual EIs (see [Table 7](#)).

For the CDEIS, responsiveness data were available in nine studies ([D'Haens 1999](#); [Hebuterne 2013](#); [Jalocha 2009](#); [Mary 1989](#); [Rutgeerts 2002](#); [Rutgeerts 2006](#); [Sipponen 2008c](#); [Sipponen 2010b](#)). Seven studies demonstrated statistically significant decreases in the CDEIS score after administration of a treatment of known efficacy ([D'Haens 1999](#); [Hebuterne 2013](#); [Rutgeerts 2002](#); [Rutgeerts 2006](#); [Sipponen 2008c](#); [Sipponen 2010b](#)). Although [Jalocha 2009](#)

showed that the CDEIS decreased after treatment, significance testing was not performed. Mary 1989 demonstrated a correlation between changes in both the CDEIS and GELS scores after a treatment of known efficacy (see Table 7). These publications suggest that the CDEIS is a responsive measure of disease activity.

Feasibility

No studies were identified that explicitly evaluated the feasibility of either the SES-CD or the CDEIS. Both instruments require scoring by endoscopist. The SES-CD is intuitively easier to use than the CDEIS, at it requires fewer calculations (see Table 8).

Methodological Quality

The COSMIN tool was used to assess the methodological quality of the included studies. In total, 20 studies assessed criterion validity for the SES-CD, CDEIS, or both indices. With regard to methodological quality, six of these studies were rated 'excellent' (Af Borjorkesten 2012; Daperno 2004; Jones 2008; Lobaton Ortega 2013; Rutgeerts 2016; Sipponen 2008a), two were rated 'good' (Rutgeerts 2010; Sipponen 2008b), five were rated 'fair' (Bamba 2012; D'Haens 1999; Ferrante 2013; Rimola 2011; Sipponen 2008c) and seven were rated 'poor' (Denis 2007; Green 2011; Molander 2013; Reinisch 2010; Sipponen 2010a; Sipponen 2010b; Vieira 2009).

Nineteen studies assessed construct validity for the SES-CD, CDEIS, or both indices. With regard to methodological quality, seven of these studies were rated 'excellent' (Af Borjorkesten 2012; Daperno 2004; Jones 2008; Khanna 2016; Mary 1989; Schoepfer 2010; Sipponen 2008a), three were rated 'good' (Rutgeerts 2002; Sipponen 2008b; Sipponen 2010b), four were rated 'fair' (D'Haens 1999; Ferrante 2013; Lobaton Ortega 2013; Sipponen 2008c), and four were rated 'poor' (Colombel 2003; Denis 2007; Ferrante 2010; Green 2011).

Five studies assessed reliability for the SES-CD, CDEIS, or both indices. With regard to methodological quality, two of these studies were rated 'excellent' (Khanna 2016; Mary 1989) and three were rated 'good' (Daperno 2004; Daperno 2014; Hebuterne 2013).

Eleven studies assessed responsiveness for the SES-CD, CDEIS, or both indices. With regard to methodological quality, three of these studies were rated 'excellent' (Mary 1989; Rutgeerts 2006; Schoepfer 2010), three were rated 'good' (Hebuterne 2013; Rutgeerts 2002; Sipponen 2010b), four were rated 'fair' (D'Haens 1999; Ferrante 2013; Sipponen 2008c; Sipponen 2010a) and one was rated 'poor' (Jalocha 2009).

None of the included studies assessed content validity or feasibility for either the SES-CD or CDEIS (See Table 9 and Table 10).

DISCUSSION

Summary of main results

Forty-five reports of 31 studies fulfilled the inclusion criteria; 43 reports of 30 studies were completed and reported results.

For the SES-CD, inter-rater reliability was assessed in four studies (Daperno 2004; Daperno 2014; Khanna 2016; Rutgeerts 2016; see Table 3). In the development study for the SES-CD (Daperno 2004), the overall ICC is 0.9815, 95% CI 0.9705 to 0.9884 and kappa for the

regions is high, however the paired raters were in the same room which introduces the potential for bias.

For the CDEIS, inter-rater reliability was assessed in six studies (Daperno 2004; Daperno 2014; Hebuterne 2013; Khanna 2016; Mary 1989; Rutgeerts 2016 see Table 3). Daperno 2014 reported the ICC for the CDEIS was 0.985 (95% CI 0.939-1.000) for average measures of video score and was 0.835 (95% CI 0.540-0.995) for single measures of video score.

With respect to validity, correlation between the CDEIS and clinical measures, including CRP, ESR, fecal calprotectin and fecal lactoferrin was also reported (see Table 6). The estimates of correlation with CRP were $r = 0.521$ (Sipponen 2010b), $r = 0.553$ (Sipponen 2008a) and $r = 0.608$ (Sipponen 2008c). Correlation between the CDEIS and CRP, ESR, fecal calprotectin and fecal lactoferrin were also explored. The corresponding values for correlation with CRP ranged from $r = 0.46$ (Jones 2008) to $r = 0.68$ (Green 2011).

For the CDEIS, responsiveness data were available in nine studies (D'Haens 1999; Hebuterne 2013; Jalocha 2009; Mary 1989; Rutgeerts 2002; Rutgeerts 2006; Sipponen 2008c; Sipponen 2010b). Seven studies demonstrated statistically significant decreases in the CDEIS score after administration of a treatment of known efficacy (D'Haens 1999; Hebuterne 2013; Rutgeerts 2002; Rutgeerts 2006; Sipponen 2008c; Sipponen 2010b). Minimal responsiveness data are available for the SES-CD. Sipponen 2010a and Sipponen 2010b demonstrated statistically significant changes in the SES-CD score after subjects were administered a treatment of known efficacy.

No studies were identified that explicitly evaluated the feasibility for either the SES-CD or the CDEIS. The SES-CD requires fewer calculations and therefore may be easier to use than the CDEIS.

Overall completeness and applicability of evidence

Although several studies have evaluated various operating properties of the SES-CD and CDEIS, many have solely focused on one or two operating properties, and contain methodological flaws. Further study is required to fully define the operating properties of these indices.

Quality of the evidence

The quality of evidence ranges from poor to excellent, and neither the SES-CD nor CDEIS has been fully validated.

Potential biases in the review process

It should be noted that this review focused on the most commonly used endoscopic indices in Crohn's disease, the CDEIS and SES-CD. However, there may be other endoscopic indices in Crohn's disease that have been subject to validation not represented in the current review.

Agreements and disagreements with other studies or reviews

We are unaware of other reviews on this topic.

AUTHORS' CONCLUSIONS

Implication for methodological research

There is an imminent need for additional validation studies to define the optimal endoscopic index for use in clinical trials of CD.

ACKNOWLEDGEMENTS

Partial funding for the Cochrane IBD Group (April 1, 2016 - March 31, 2018) has been provided by Crohn's and Colitis Canada (CCC).

REFERENCES

References to studies included in this review

Af Bjorkesten 2012 {published data only}

Af Bjorkesten CG, Nieminen U, Arkkila P, Turunen U, Sipponen T, Farkkila M. Endoscopy at three months from start of anti-TNF as a predictor for long-term endoscopic remission in Crohn's disease. *Journal of Crohn's and Colitis* 2012;**6**:S162.

* Af Bjorkesten CG, Nieminen U, Turunen U, Arkkila P, Sipponen T, Farkkila M. Surrogate markers and clinical indices, alone or combined, as indicators for endoscopic remission in anti-TNF-treated luminal Crohn's disease. *Scandinavian Journal of Gastroenterology* 2012;**47**(5):528-37.

Bamba 2012 {published data only}

Bamba S, Andoh A, Ban H, Imaeda H, Mochizuki Y, Inatomi O. Assessment and validation of endoscopic activity indices and their correlation to blood biochemical markers in Crohn's disease. *Journal of Gastroenterology and Hepatology* 2012;**27**:398.

Colombel 2003 {published data only}

Colombel JF, Lemann M, Bouhnik Y, Duclos B, Soule JC, Lerebours E, et al. Endoscopic healing of Crohn's ileo-colitis with azathioprine. *Gastroenterology* 2003;**124**(4 Suppl 1):A196-7.

D'Haens 1999 {published data only}

D'Haens G, Van Deventer S, Van Hogezand R, Chalmers D, Kothe C, Baert F, et al. Endoscopic and histological healing with infliximab anti-tumor necrosis factor antibodies in Crohn's disease: a European multicenter trial. *Gastroenterology* 1999;**116**(5):1029-34.

Daperno 2004 {published data only}

Daperno M, D'Haens G, Van Assche G, Baert F, Bulois P, Maunoury V, et al. Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. *Gastrointestinal Endoscopy* 2004;**60**(4):505-12.

Daperno 2014 {published data only}

Daperno M, Comberlato M, Bossa F, Biancone L, Bonanomi A, Cassinotti A, et al. Concordance in IBD endoscopic scoring requires expertise and training: preliminary results of an ongoing IG-IBD study. *Journal of Crohn's and Colitis* 2012;**6**:S125.

Daperno M, Comberlato M, Bossa F, Biancone L, Bonanomi A, Cassinotti A, et al. Interobserver agreement in IBD endoscopic scoring: preliminary results of an ongoing IG-IBD study. *Gastroenterology* 2010;**142**(5 Suppl 1):S536.

Daperno M, Comberlato M, Bossa F, Biancone L, Bonanomi A, Cassinotti A, et al. Interobserver agreement in IBD scores requires expertise and education: preliminary results from an ongoing IG-IBD study. *Digestive and Liver Disease* 2012;**44**:S80-1.

* Daperno M, Comberlato M, Bossa F, Biancone L, Bonanomi AG, Cassinotti A, et al. Inter-observer agreement in endoscopic scoring systems: preliminary report of an ongoing

study from the Italian Group for Inflammatory Bowel Disease (IG-IBD). *Digestive and Liver Disease* 2014;**46**(11):969-73.

Daperno M, Comberlato M, Bossa F, Bonanomi A, Lombardi G, Biancone L, et al. Increasing interobserver agreement on IBD endoscopic scoring systems: results from the IGIBDendo educational program. *Digestive and Liver Disease* 2014;**46**:S4.

Daperno M, Comberlato M, Bossa F, Bonanomi AG, Lombardi G, Biancone L, et al. Increasing interobserver agreement on ibd endoscopic scoring systems: results from the IGIBDendo educational program. *Gastroenterology* 2014;**146**(5 Suppl 1):S-234.

Denis 2007 {published data only}

Denis MA, Reenaers C, Fontaine F, Belaiche J, Louis E. Assessment of endoscopic activity index and biological inflammatory markers in clinically active Crohn's disease with normal C-reactive protein serum level. *Inflammatory Bowel Diseases* 2007;**13**(9):1100-5.

Ferrante 2010 {published data only}

Ferrante M, Noman M, Vermeire S, Van Assche GA, Rutgeerts PJ. Evolution of endoscopic activity scores under placebo therapy in Crohn's disease. *Gastroenterology* 2010;**138**(5 Suppl 1):S358.

Ferrante 2013 {published data only}

Ferrante M, Colombel JF, Sandborn WJ, Reinisch W, Mantzaris GJ, Kornbluth A. Evolution of endoscopic activity scores in patients with Crohn's disease under azathioprine and/or infliximab: a post-hoc analysis of the sonic data. *Gastroenterology* 2011;**140**(5 Suppl 1):S422.

* Ferrante M, Colombel JF, Sandborn WJ, Reinisch W, Mantzaris GJ, Kornbluth A, et al. Validation of endoscopic activity scores in patients with Crohn's disease based on a post hoc analysis of data from SONIC. *Gastroenterology* 2013;**145**(5):978-86.

Green 2011 {published data only}

Green EH, Al-Hawary MM, Punglia D, Zimmermann EM, Adler J. Endoscopic score extracted from scope report correlates with Crohn's disease activity and CT enterography findings. *Gastroenterology* 2011;**140**(5 Suppl 1):S426.

Hebuterne 2013 {published data only}

* Hebuterne X, Lemann M, Bouhnik Y, Dewit O, Dupas JL, Mross M, et al. Endoscopic improvement of mucosal lesions in patients with moderate to severe ileocolonic Crohn's disease following treatment with certolizumab pegol. *Gut* 2013;**62**(2):201-8.

Hebuterne X, Lemann M, Louis Dupas J, Bouhnik Y, Allez M, Louis E, et al. Endoscopic improvement in patients with active Crohn's disease treated with certolizumab pegol: results of blinded central reading of recorded endoscopies from the MUSIC study. *American Journal of Gastroenterology* 2010;**105**:S418-9.

Jalocha 2009 {published data only}

Jalocha L, Wojtun S, Gil J, Dyrła P. Efficacy of anti-TNF-alpha agents in mucosal healing in Crohn disease. *Przegląd Gastroenterologiczny* 2009;**4**(4):184-7.

Jones 2008 {published data only}

Jones J, Loftus EV Jr, Panaccione R, Chen LS, Peterson S, McConnell J, et al. Relationships between disease activity and serum and fecal biomarkers in patients with Crohn's disease. *Clinical Gastroenterology and Hepatology* 2008;**6**(11):1218-24.

Khanna 2016 {published data only}

Khanna R, D'Haens G, Rutgeerts P, McDonald J, Daperno M, Feagan B, et al. Reliability of central readers in the evaluation of endoscopic disease activity in Crohn's disease. *Inflammatory Bowel Diseases* 2013;**19**:s50.

Khanna R, Zou G, D'Haens G, Rutgeerts P, McDonald JW, Daperno M, et al. Agreement among central readers in the evaluation of endoscopic disease activity in Crohn's disease. *Journal of Crohn's and Colitis* 2014;**8**:S13-4.

* Khanna R, Zou G, D'Haens G, Rutgeerts P, McDonald JW, Daperno M, et al. Reliability among central readers in the evaluation of endoscopic findings from patients with Crohn's disease. *Gut* 2016;**65**(7):1119-25.

Khanna R, Zou G, D'Haens GR, Rutgeerts PJ, McDonald JW, Daperno M, et al. Agreement among central readers in the evaluation of endoscopic disease activity in Crohn's disease. *Gastroenterology* 2014;**146**(5 Suppl 1):S-175.

Lobaton Ortega 2013 {published data only}

* Lobaton Ortega T, Lopez-Garcia A, Rodriguez-Moranta F, Ruiz A, Rodriguez L, Guardiola J. A new rapid test for fecal calprotectin predicts endoscopic remission and postoperative recurrence in Crohn's disease. *Journal of Crohn's and Colitis* 2013;**7**(12):e641-51.

Lobaton Ortega T, Rodriguez-Moranta F, Guardiola J, Sanchez-Pastor E, Rodriguez Alonso L, Ruiz-Cerulla S, et al. A new rapid test for fecal calprotectin (FC) predicts mucosa healing in Crohn's disease (CD). *Journal of Crohn's and Colitis* 2012;**6**:S81.

Mary 1989 {published data only}

Mary JY, Modigliani R. Development and validation of an endoscopic index of the severity for Crohn's disease: a prospective multicentre study. Groupe d'Etudes Thérapeutiques des Affections Inflammatoires du Tube Digestif (GETAID). *Gut* 1989;**30**(7):983-9.

Molander 2013 {published data only}

Molander P, Sipponen T, Kemppainen H, Jussila A, Blomster T, Koskela R, et al. Achievement of deep remission during scheduled maintenance therapy with TNFalpha-blocking agents in IBD. *Journal of Crohn's and Colitis* 2013;**7**(9):730-5.

Reinisch 2010 {published data only}

Reinisch W, Rutgeerts PJ, Panaccione R, D'Haens GR, Thakkar R, Yu AP, et al. Prediction of long-term clinical remission for adalimumab-treated patients with Crohn's disease by

identification of appropriate dichotomizing points for SES-CD. *Gastroenterology* 2010;**138**(5 Suppl 1):S8.

Rimola 2011 {published data only}

Rimola J, Ordas I, Rodriguez S, Garcia-Bosch O, Aceituno M, Llach J, et al. Magnetic resonance imaging for evaluation of Crohn's disease: validation of parameters of severity and quantitative index of activity. *Inflammatory Bowel Diseases* 2011;**17**(8):1759-68.

Rutgeerts 2002 {published data only}

Rutgeerts P, Van Assche G, Van Deventer S, Bao W, Keenan G, Olson A, et al. Infliximab maintenance treatment strategy results in mucosal healing in patients with Crohn's disease. *Gastroenterology* 2002;**122**(Suppl 4):A618.

Rutgeerts 2006 {published data only}

Rutgeerts P, Diamond RH, Bala M, Olson A, Lichtenstein GR, Bao W, et al. Scheduled maintenance treatment with infliximab is superior to episodic treatment for the healing of mucosal ulceration associated with Crohn's disease. *Gastrointestinal Endoscopy* 2006;**63**(3):433-42.

Rutgeerts 2010 {published data only}

Rutgeerts PJ, Reinisch W, Thakkar R, Wu EQ, Kaltenboeck A, Yang M, et al. Early mucosal healing status predicts long-term clinical benefits for adalimumab-treated patients with moderate to severe Crohn's disease. *Gastroenterology* 2010;**138**(5 Suppl 1):S85.

Rutgeerts PJ, Thakkar R, Wu EQ, Kaltenboeck A, Yang M, Mulani P, et al. Quality-of-life benefits for adalimumab-treated patients with moderate to severe Crohn's disease are predicted by early mucosal healing status. *Gastroenterology* 2010;**138**(5 Suppl 1):S320.

Rutgeerts 2016 {published data only}

Rutgeerts P, Reinisch W, Colombel JF, Sandborn WJ, D'Haens G, Petersson J, et al. Agreement of site and central readings of ileocolonoscopy scores in Crohn's disease: comparison using data from the EXTEND trial. *Gastrointestinal Endoscopy* 2016;**83**(1):188-97.

Schoepfer 2010 {published data only}

Schoepfer AM, Beglinger C, Straumann A, Trummler M, Vavricka SR, Bruegger LE, et al. Fecal calprotectin correlates more closely with the Simple Endoscopic Score for Crohn's disease (SES-CD) than CRP, blood leukocytes, and the CDAl. *American Journal of Gastroenterology* 2010;**105**(1):162-9.

Sipponen 2008a {published data only}

Sipponen T, Savilahti E, Kolho KL, Nuutinen H, Turunen U, Farkkila M. Crohn's disease activity assessed by fecal calprotectin and lactoferrin: correlation with Crohn's disease activity index and endoscopic findings. *Inflammatory Bowel Diseases* 2008;**14**(1):40-6.

Sipponen 2008b {published data only}

Sipponen T, Karkkainen P, Savilahti E, Kolho KL, Nuutinen H, Turunen U, et al. Correlation of faecal calprotectin and lactoferrin with an endoscopic score for Crohn's disease

and histological findings. *Alimentary Pharmacology and Therapeutics* 2008;**28**(10):1221-9.

Sipponen 2008c {published data only}

Sipponen T, Savilahti E, Karkkainen P, Kolho KL, Nuutinen H, Turunen U, et al. Fecal calprotectin, lactoferrin, and endoscopic disease activity in monitoring anti-TNF-alpha therapy for Crohn's disease. *Inflammatory Bowel Diseases* 2008;**14**(10):1392-8.

Sipponen 2010a {published data only}

Sipponen T, Bjorkestén CG, Farkkila M, Nuutinen H, Savilahti E, Kolho KL. Faecal calprotectin and lactoferrin are reliable surrogate markers of endoscopic response during Crohn's disease treatment. *Scandinavian Journal of Gastroenterology* 2010;**45**(3):325-31.

Sipponen 2010b {published data only}

Sipponen T, Nuutinen H, Turunen U, Farkkila M. Endoscopic evaluation of Crohn's disease activity: comparison of the CDEIS and the SES-CD. *Inflammatory Bowel Diseases* 2010;**16**(12):2131-6.

Vieira 2009 {published data only}

Vieira A, Fang CB, Rolim EG, Klug WA, Steinwurz F, Rossini LG, et al. Inflammatory bowel disease activity assessed by fecal calprotectin and lactoferrin: correlation with laboratory parameters, clinical, endoscopic and histological indexes. *BMC Research Notes* 2009;**2**:221.

References to studies excluded from this review

Bondjemah 2012 {published data only}

Bondjemah V, Mary JY, Jones J, Sandborn W, Schoepfer A, Louis E, et al. Fecal calprotectin and CRP as biomarkers of endoscopic activity in Crohn's disease: a meta-study. *Journal of Crohn's and Colitis* 2012;**6**(1):S63.

Mantzaris 2009 {published data only}

Mantzaris GJ, Christidou A, Sfakianakis M, Roussos A, Koilakou S, Petraki K, et al. Azathioprine is superior to budesonide in achieving and maintaining mucosal healing and histologic remission in steroid-dependent Crohn's disease. *Inflammatory Bowel Diseases* 2009;**15**(3):375-82.

Modigliani 1987 {published data only}

Modigliani R, Mary JY. Reproducibility of colonoscopic findings in Crohn's disease: A prospective multicenter study of interobserver variation. *Digestive Diseases and Sciences* 1987;**32**(12):1370-9.

Sandborn 2012a {published data only}

Sandborn W, Colombel JF, Allez M, Dupas JL, Dewit O, D'Haens G, et al. Certolizumab pegol plasma concentrations and endoscopic and clinical outcomes in Crohn's disease: a post hoc analysis of the MUSIC trial. *American Journal of Gastroenterology* 2012;**107** Suppl 1:S674-5.

Sandborn 2012b {published data only}

Sandborn W, Panaccione R, Colombel JF, Louis E, Yang M, Thakkar R, et al. Development of three practical indices for

mucosal healing among patients with moderate to severe Crohn's disease. *Journal of Crohn's and Colitis* 2012;**6**(1):S3.

Sandborn WJ, Panaccione R, Colombel JF, Louis E, Yang M, Thakkar R, et al. Development of three practical indices for mucosal healing among patients with moderate to severe Crohn's disease. *Gastroenterology* 2012;**142**(5 Suppl 1):S537.

Sandborn WJ, Panaccione R, Colombel JF, Louis E, Yang M, Thakkar R, et al. Development of three practical indices for mucosal healing among patients with moderate to severe Crohn's disease. *Inflammatory Bowel Diseases* 2012;**18**:S39-40.

Sandborn 2012c {published data only}

Sandborn W, Colombel JF, Louis E, Panaccione R, Yang M, Chao J, et al. Adalimumab sustains deep remission for 3 years: Data from CHARM and ADHERE. *Journal of Crohn's and Colitis* 2012;**6**(1):S169-70.

Sandborn W, Colombel JF, Louis E, Panaccione R, Yang M, Chao J, et al. Adalimumab sustains deep remission for 3 years: data from CHARM and ADHERE. *Inflammatory Bowel Diseases* 2012;**18**:S38.

Ziech 2012 {published data only}

Ziech ML, Lavini C, Caan MW, Nio CY, Stokkers PC, Bipat S, et al. Dynamic contrast-enhanced MRI in patients with luminal Crohn's disease. *European Journal of Radiology* 2012;**81**(11):3019-27.

References to studies awaiting assessment

Rutgeerts 2004 {published data only}

References to ongoing studies

Buisson 2015 {published data only}

Buisson A, Filippi J, Amiot A, Cadiot G, Allez M, Marteau PR, et al. Definitions of the endoscopic lesions in Crohn's disease: reproducibility study and GETAID expert consensus. *Gastroenterology* 2015;**148**(4 Suppl 1):S445-6.

Buisson A, Filippi J, Amiot A, Cadiot G, Allez M, Marteau PR, et al. Definitions of the endoscopic lesions in Crohn's disease: reproducibility study and GETAID expert consensus. *Journal of Crohn's and Colitis* 2015;**9**:S175-6.

Additional references

Cellier 1994

Cellier C, Sahmoud T, Froguel E, Adenis A, Belaiche J, Bretagne JF, et al. Correlations between clinical activity, endoscopic severity, and biological parameters in colonic or ileocolonic Crohn's disease. A prospective multicentre study of 121 cases. *Gut* 1994;**35**:231-5.

Deyo 1991

Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and

strategies for evaluation. *Controlled Clinical Trials* 1991;**12**(4 Suppl):142S-158S.

FDA 2013

U.S. Food and Drug Administration (FDA) Workshop: Gastroenterology regulatory endpoints and the advancement of therapeutics (transcript). <https://www.regulations.gov/#!documentDetail;D=FDA-2013-N-0001-0130> (Accessed March 30, 2016).

Feagan 2013

Feagan BG, Sandborn WJ, D'Haens G, Pola S, McDonald JW, Rutgeerts P, et al. The role of centralized reading of endoscopy in a randomized controlled trial of mesalamine for ulcerative colitis. *Gastroenterology* 2013;**145**(1):149-57.

Kirshner 1985

Kirshner B, Guyatt G. A methodological framework for assessing health indices. *Journal of Chronic Diseases* 1985;**38**(1):27-36.

Regueiro 2011

Regueiro M, Kip KE, Schraut W, Baidoo L, Sepulveda AR, et al. Crohn's disease activity index does not correlate with endoscopic recurrence one year after ileocolonic resection. *Inflammatory Bowel Diseases* 2011;**17**(1):118-26.

Zou 2005

Zou GY. Quantifying responsiveness of quality of life measures without an external criterion. *Quality of Life Research* 2005;**14**(6):1545-52.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Af Borjorkesten 2012

Methods	64 patients 210 supervised colonoscopies Each SES-CD scored by at least 2 gastroenterologists
Data	Patient characteristics: Anti-TNF treated luminal CD Mean age: 38.8 (18-69) Females/Males: 103/107
Comparisons	SES-CD evaluated
Outcomes	See additional tables
Blinding	Partially blinded (fecal calprotectin but not to other lab or clinical data)
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	No	Readers blinded to fecal calprotectin but not to other lab or clinical data
Independent observation? All outcomes	Unclear	Not adequately described

Bamba 2012

Methods	Retrospective chart review
---------	----------------------------

Bamba 2012 (Continued)

158 patients with balloon enteroscopy

Data	Not described
Comparisons	SES-CD evaluated
Outcomes	See additional tables
Blinding	No
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	No	It does not appear that raters were blinded
Independent observation? All outcomes	Unclear	Not adequately described

Colombel 2003

Methods	45 patients randomized to azathioprine or placebo
Data	Patient characteristics: Female/Male: 24/21 Mean age: 37
Comparisons	CDEIS evaluated
Outcomes	See additional tables
Blinding	Not adequately described
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation? All outcomes	Unclear	Not adequately described

D'Haens 1999

Methods	30 patients with pre- and post-treatment colonoscopies recorded
---------	---

D'Haens 1999 (Continued)

8 patients received placebo, 22 patients received infliximab

Data	<p>Patient characteristics:</p> <p>Female/Male: 18/12</p> <p>Mean Age Placebo: 34.4</p> <p>Mean Age IFX: 31.4</p>
Comparisons	CDEIS evaluated
Outcomes	See additional tables
Blinding	No
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	No	Local endoscopist performed scoring; not blinded to clinical information
Independent observation? All outcomes	Unclear	Not applicable - single review

Daperno 2004

Methods	<p>191 patients:</p> <p>Development of SES-CD: 70 patients (endoscopist paired with observer for 35 of 70)</p> <p>Validation of SES-CD: 121 patients (endoscopist paired with second observer for 36 of 121)</p> <p>Endoscopist scored directly during procedure (70 +121)</p> <p>A second observer scored the colonoscopy from a television monitor (35 + 36) for inter-rater variation study</p>
Data	<p>Patient characteristics:</p> <p>Age <40: 169/191</p> <p>Age ≥40: 22/191</p> <p>Female:Male = 123/68</p> <p>Median (range) disease duration: 9 yrs (0.5-37)</p> <p>CDAI median range: 227 (-6 to 573)</p>
Comparisons	SES-CD and CDEIS evaluated
Outcomes	See additional tables
Blinding	No
Notes	

Risk of bias

Item	Authors' judgement	Description
------	--------------------	-------------

Endoscopic scoring indices for evaluation of disease activity in Crohn's disease (Review)

Daperno 2004 (Continued)

Blinding?	No	Local endoscopist performed scoring; not blinded to clinical information
Independent observation? All outcomes	Yes	The endoscopist and the observer did not communicate during the colonoscopic examination, and they completed the endoscopic scoring sheets independently

Daperno 2014

Methods	8 endoscopic videos scored
Data	Patient characteristics: Patients with luminal CD
Comparisons	SES-CD and CDEIS evaluated
Outcomes	See additional tables
Blinding	Unclear
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	Unclear	It is unclear whether the endoscopist and observer were blinded to clinical information
Independent observation? All outcomes	Yes	The endoscopist and the observer did not communicate during the colonoscopic examination, and they completed the endoscopic scoring sheets independently

Denis 2007

Methods	28 patients with clinically active CD (CDAI > 150 with normal serum CRP $\leq$ 6 mg/L) 28 colonoscopies
Data	Patient characteristics: Female/Male: 19/9 Median age: 46 (23-68)
Comparisons	CDEIS evaluated
Outcomes	See additional tables
Blinding	Unclear
Notes	

Risk of bias
Endoscopic scoring indices for evaluation of disease activity in Crohn's disease (Review)

Denis 2007 (Continued)

Item	Authors' judgement	Description
Blinding?	Unclear	It is unclear whether clinical information was concealed when the CDEIS was calculated
Independent observation? All outcomes	Unclear	Not applicable - single review

Ferrante 2010

Methods	24 patients 48 videos (24 pre/24 post)	
Data	Patient characteristics: All patients had moderate to severe CD	
Comparisons	SES-CD and CDEIS evaluated	
Outcomes	See additional tables	
Blinding	Yes	
Notes		

Risk of bias

Item	Authors' judgement	Description
Blinding?	Yes	The central reader was blinded to study number, treatment and recording sequence
Independent observation? All outcomes	Unclear	Not applicable - single review

Ferrante 2013

Methods	172 patients (120 with CRP elevated at baseline)	
Data	Patient characteristics: Female/Male: 53%/47% Median age: 34.0 years	
Comparisons	SES-CD and CDEIS evaluated	
Outcomes	See additional tables	
Blinding	Yes	

Ferrante 2013 (Continued)

Notes

Risk of bias

Item	Authors' judgement	Description
Blinding?	Yes	The central reader who calculated the SES-CD and CDEIS based on videos and was unaware of study group assignment
Independent observation? All outcomes	Yes	Not applicable - single review

Green 2011

Methods	18 patients Performed chart review of colonoscopy (methods unclear)
Data	Patient characteristics: Age (20-58 years)
Comparisons	SES-CD evaluated
Outcomes	See additional tables
Blinding	Yes
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	Yes	Review was performed in a blinded manner
Independent observation? All outcomes	Unclear	Not adequately described

Hebuterne 2013

Methods	89 patients (44 with paired endoscopy) 88 videos 1 site reader and 4 central readers 2 pairs of 2 readers split reading (only used central readers for analysis) 78 completed endoscopy at Week 10 53 completed endoscopy at Week 54 133 videos with consent (49 at baseline, 51 at week 10, 33 at week 54) Each read by a pair of central readers (266 reads)
---------	---

Endoscopic scoring indices for evaluation of disease activity in Crohn's disease (Review)

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Hebuterne 2013 (Continued)

Data	Patient characteristics: Active endoscopic disease (ulcerations in ≥ 2 segments with CDEIS score ≥ 8) Females/Males: 59/30 Age mean: 30.2 +/- 9.9	
Comparisons	CDEIS evaluated	
Outcomes	See additional tables	
Blinding	Yes	
Notes		
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Blinding?	Yes	Expert endoscopists were blinded
Independent observation? All outcomes	Yes	Assumed yes

Jalocha 2009

Methods	40 scopes (20 pre, 20 post) One local endoscopist scored and same endoscopist re-evaluated after 10 weeks of infliximab or 12 weeks of adalimumab treatment	
Data	Patient characteristics: 20 patients (10 infliximab, 10 adalimumab)	
Comparisons	CDEIS evaluated	
Outcomes	See additional tables	
Blinding	No	
Notes		
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation? All outcomes	Unclear	Not applicable - single review

Jones 2008

Methods	164 patients
Data	Patient characteristics: Female/Male: 97/68 Mean age: 41
Comparisons	SES-CD evaluated in 157 patients
Outcomes	See additional tables
Blinding	Yes
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	Unclear	Endoscopic disease activity was measured by endoscopists who were blinded to the results of the laboratory studies (unclear whether they were blinded to clinical information)
Independent observation? All outcomes	Unclear	"Repeat review of the SES-CD scoring sheets"

Khanna 2016

Methods	Videos of colonoscopies from 50 patients were viewed in random order by four central readers Data were used to assess intra-rater and inter-rater reliability for CDEIS, SES-CD and a global evaluation of lesion severity (GELS) Central readers participated in a consensus process that identified common sources of disagreement
Data	Patient characteristics: CDAI score of 220 and a CRP concentration ≥ 5 mg/L or a faecal calprotectin concentration ≥ 250 mg/g
Comparisons	CDEIS and SES-CD evaluated
Outcomes	Intra-rater and inter-rater reliability See additional tables
Blinding	Yes
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	Yes	Central readers were blind to all study information

Khanna 2016 (Continued)

Independent observation? All outcomes	Yes	Central readers independently reviewed recordings
--	-----	---

Lobaton Ortega 2013

Methods	46 patients with colonoscopy in original abstract 89 complete colonoscopies in 2013 follow-up publication	
Data	Patient characteristics (from 2013 follow-up publication): Female/Male: 65/50 Age at diagnosis: Age ≤16 (3%), Age 17-40 (90%), Age > 40 (7%) CD duration: 8 years (IQR 5-21) CDAI: 50.8 (0-120)	
Comparisons	SES-CD and CDEIS evaluated	
Outcomes	See additional tables	
Blinding	Yes	
Notes		

Risk of bias

Item	Authors' judgement	Description
Blinding?	Yes	Both endoscopists were blind to all study variables
Independent observation? All outcomes	Unclear	Not adequately described

Mary 1989

Methods	CDEIS construction and validation study	
Data	Patient characteristics: 66% of the reproducibility set and 62% of the responsiveness set had active disease (CDAI >150)	
Comparisons	CDEIS evaluated	
Outcomes	See additional data tables	
Blinding	Unclear	
Notes	Original CDEIS validation study	

Risk of bias

Mary 1989 (Continued)

Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation? All outcomes	Yes	The endoscopist and the observer did not communicate during the colonoscopic examination, and they completed the endoscopic scoring sheets independently

Molander 2013

Methods	183 CD patients (109 Infliximab, 74 adalimumab)	
Data	Patient characteristics: Age Median: 32(12-69) Female/Male: 85/98 Disease duration: 8(0-33)	
Comparisons	SES-CD evaluated	
Outcomes	See additional tables	
Blinding	No	
Notes		

Risk of bias

Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation? All outcomes	Unclear	Not applicable - single review

Reinisch 2010

Methods	186 patients with baseline, week 12 and week 24 results receiving adalimumab	
Data	Patient characteristics: Moderate to severe ileocolonic CD	
Comparisons	SES-CD evaluated	
Outcomes	See additional tables	
Blinding	Unclear	
Notes		

Risk of bias
Endoscopic scoring indices for evaluation of disease activity in Crohn's disease (Review)

Reinisch 2010 (Continued)

Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation? All outcomes	Unclear	Not applicable - single review

Rimola 2011

Methods	48 patients with colonoscopy Single review with two readers: one performed examination and other recorded findings
Data	Patient characteristics: Female/Male: 33/15 Age at diagnosis: 25.5 (18.3-34.9) CD duration: 6.9 (1-10.6)
Comparisons	CDEIS evaluated
Outcomes	See additional tables
Blinding	Yes
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	Yes	Investigators performing ileocolonoscopy or MRI were unaware of clinical data
Independent observation? All outcomes	Unclear	Not applicable - single review

Rutgeerts 2002

Methods	66 patients (24 placebo, 19 infliximab 5mg/kg, 23 infliximab 10 mg/kg) Endoscopy at baseline, week 10, week 54 (198 in total)
Data	Patient characteristics: CDAI 220-400
Comparisons	CDEIS evaluated
Outcomes	See additional tables
Blinding	Unclear

Endoscopic scoring indices for evaluation of disease activity in Crohn's disease (Review)

Rutgeerts 2002 (Continued)

Notes

Risk of bias

Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation? All outcomes	Unclear	Not applicable - single review

Rutgeerts 2006

Methods	Endoscopic evaluation was performed by the endoscopist at the local facility, and a video recording was assessed by a central reviewer Correlation between CDAI and CDEIS evaluated
Data	Patient characteristics: Eligible patients were ≥ 18 years of age with Crohn's disease of at least 3 months' duration (colitis, ileitis, or ileocolitis confirmed by radiography or endoscopy) and a CDAI score of 220-400
Comparisons	CDEIS evaluated
Outcomes	See additional tables
Blinding	Yes
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	Yes	The central reader was blinded to all study information
Independent observation? All outcomes	Unclear	Not applicable - single review

Rutgeerts 2010

Methods	Patients underwent an endoscopic assessment of mucosal activity at baseline, week 12, and week 52 Post-hoc analysis of 62 patients randomized to adalimumab who also were not missing endoscopic assessment scores at week 12 Week 52 CDAI scores and clinical remission status were compared between patients with and without mucosal healing at week 12
Data	Patient characteristics:

Rutgeerts 2010 (Continued)

Not described

Comparisons	SES-CD evaluated
Outcomes	See additional tables
Blinding	Unclear
Notes	Abstract publication

Risk of bias

Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation? All outcomes	Unclear	Not adequately described

Rutgeerts 2016

Methods	Agreement between readers for CDEIS–scored endoscopies and SES-CD–scored endoscopies from 19 sites in the EXTEND study was evaluated at baseline and weeks 12 and 52	
Data	Patient characteristics: Adult patients with moderately to severely active CD, ileocolonic CD for 4 months, CDAI 220-450, and mucosal ulceration documented by recorded ileocolonoscopy at screening	
Comparisons	CDEIS and SES-CD	
Outcomes	See additional tables	
Blinding	Yes	
Notes		

Risk of bias

Item	Authors' judgement	Description
Blinding?	Yes	Blinded to local reading
Independent observation? All outcomes	Yes	Assumed yes, the central reader reviewed videos

Schoepfer 2010

Methods	122 CD patients, 43 healthy controls 140 colonoscopies in 122 CD patients	
---------	--	--

Schoepfer 2010 (Continued)

Data	Patient characteristics: Female/Male: 71/51 Mean age: 42 +/- 16 (19-83) Mean disease duration: 13.2 +/- 6.1	
Comparisons	SES-CD evaluated	
Outcomes	See additional tables	
Blinding	Yes	
Notes		
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Blinding?	Yes	CDAI was measured by a physician not performing the colonoscopy Those performing the endoscopies were unaware of the results of the CDAI, fecal calprotectin, CRP and blood leukocytes
Independent observation? All outcomes	Unclear	Not applicable - single review

Sipponen 2008a

Methods	77 CD patients 106 colonoscopies	
Data	Patient characteristics: Female/Male: 38/39 Median age: 33.5 (19-70) Mean disease duration: 9.2	
Comparisons	CDEIS evaluated	
Outcomes	See additional tables	
Blinding	Unclear	
Notes		
<i>Risk of bias</i>		
Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation?	Unclear	Not applicable - single review

Sipponen 2008a (Continued)

All outcomes

Sipponen 2008b

Methods	61 patients 87 ileocolonoscopies (37 scoped once, 22 twice, 2 three times)
Data	Patient characteristics: Age median (range): 33 (19-70) Female/Male: 31/30 Disease type (inflammatory/stenosing/penetrating): 31/23/7
Comparisons	SES-CD evaluated
Outcomes	See additional tables
Blinding	Unclear
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	Unclear	The gastroenterologists who scored using the SES-CD were unaware of faecal assay results, however it is unclear whether they were unaware of CDAI scores
Independent observation? All outcomes	Unclear	Not applicable - single review

Sipponen 2008c

Methods	15 patients (14 infliximab, 1 adalimumab) 30 colonoscopies (baseline and week 12, or week 10 for adalimumab patients)
Data	Patient characteristics: Female/Male: 6/9 Median age: 25 (19-44) Median disease duration: 5.1 (0.4-27)
Comparisons	CDEIS evaluated
Outcomes	See additional tables
Blinding	Unclear
Notes	

Risk of bias

Sipponen 2008c (Continued)

Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation? All outcomes	Yes	Not applicable - single review

Sipponen 2010a

Methods	19 patients 38 colonoscopies (19 at baseline and 19 at 4-6 months)
Data	Patient characteristics: Female/male: 9/10 Mean age 31.2 (19-51) Mean disease duration: 6.8 (0-27)
Comparisons	SES-CD evaluated
Outcomes	See additional tables
Blinding	Unclear
Notes	

Risk of bias

Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation? All outcomes	Unclear	Not applicable - single review

Sipponen 2010b

Methods	86 patients, 32 with second endoscopy
Data	Patient characteristics: Female/male: 43/43 Median age: 35 (19-70) Median disease duration: 6.9 (0-31.1)
Comparisons	SES-CD and CDEIS evaluated
Outcomes	See additional tables
Blinding	Unclear

Endoscopic scoring indices for evaluation of disease activity in Crohn's disease (Review)

Sipponen 2010b (Continued)

Notes

Risk of bias

Item	Authors' judgement	Description
Blinding?	Unclear	Not adequately described
Independent observation? All outcomes	Unclear	Not applicable - single review

Vieira 2009

Methods	38 patients with double balloon enteroscopy 2 endoscopists reviewed videos	
Data	Patient characteristics: Female/Male: 24/14 Median age: 37 (18-64)	
Comparisons	CDEIS	
Outcomes	See additional tables	
Blinding	Yes	
Notes		

Risk of bias

Item	Authors' judgement	Description
Blinding?	Yes	Endoscopists were unaware of the other test results
Independent observation? All outcomes	Unclear	Not applicable - single review

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Bondjemah 2012	Pooled analysis of 6 studies
Mantzaris 2009	Not a validation study
Modigliani 1987	Did not utilize the SES-CD or CDEIS: Endoscopist Pair Method (EPM)
Sandborn 2012a	Not a validation study

Endoscopic scoring indices for evaluation of disease activity in Crohn's disease (Review)

Study	Reason for exclusion
Sandborn 2012b	Not a validation study: attempt to develop a proxy measure for identifying patients who are likely to have achieved mucosal healing in Crohn's disease using clinical trial data
Sandborn 2012c	Not a validation study: uses same proxy measure utilized in Sandborn 2012b to identify patients who achieved 'deep remission'
Ziech 2012	Not a validation study: dynamic contrast-enhanced MRI

Characteristics of studies awaiting assessment *[ordered by study ID]*

[Rutgeerts 2004](#)

Methods	53 patients had video-recorded colonoscopies performed at weeks 0 and 10
Data	Patient characteristics: Patients had active disease and participated in the ENACT-1 trial
Comparisons	CDEIS
Outcomes	Mean change in CDEIS score; proportion of patients who were ulcer-free at week 10
Notes	

Characteristics of ongoing studies *[ordered by study ID]*

[Buisson 2015](#)

Trial name or title	Definitions of the endoscopic lesions in Crohn's disease: reproducibility study and GETAID expert consensus
Methods	This expert consensus is being performed according to the Delphi process The reproducibility study assesses the intra and inter-observer agreement
Data	Results from the expert consensus have been reported
Comparisons	CDEIS
Outcomes	Definitions of aphtoid erosions (AE), superficial ulcerations (SU), deep ulcerations (DU), stenosis and fistulas Intra-rater and inter-rater reliability
Starting date	Not described
Contact information	A. Buisson, University Hospital Estaing, Gastroenterology Department, Clermont-Ferrand, France
Notes	

ADDITIONAL TABLES

Table 1. Summary of Included Studies

Study ID	Index	No. of Readers	No. of CD Patients	No. of Endoscopies and Reviews	Material Reviewed	Independent Review	Blinded Review	Operating Properties Assessed
Af Bor-jorkesten 2012	SES-CD	≥ 1	64	210 colonoscopies Each SES-CD scored by at least 2 gastroenterologists	Supervised colonoscopies	?	No Fecal calprotectin blinded but not other lab or clinical data	Criterion validity; construct validity
Bamba 2012	SES-CD	?	158	158	Balloon enteroscopy, from medical charts	?	No	Criterion validity; construct validity
Colombel 2003	CDEIS	?	45 (AZA or placebo)	45	Colonoscopy	?	?	Construct validity
D'Haens 1999	CDEIS	≥ 1	30 (Placebo: 8, IFX: 22)	60 (30 pre, 30 post)	Colonoscopy (also recorded)	N/A- single review	No	Criterion validity; construct validity; responsiveness
Daperno 2004	SES-CD CDEIS	≥ 1	191	a) Development of SES-CD: 70 patients (endoscopist paired with observer for 35 of 70) b) Validation of SES-CD: 121 patients (endoscopist paired with second observer for 36 of 121)	Colonoscopy and live video feed: local endoscopist (70 +121), second observer from a television monitor (35 + 36)	Yes	No	Reliability; criterion validity; construct validity
Daperno 2014	SES-CD	13	8	8	Videos	Yes	?	Reliability

Table 1. Summary of Included Studies (Continued)
CDEIS

Denis 2007	CDEIS	1	28	28	Colonoscopy	N/A - single review	?	Criterion validity; construct validity
Ferrante 2010	SES-CD	?	24	48 videos (24 pre, 24 post)	Videos	N/A - single review	Yes	Construct validity
Ferrante 2013	SES-CD	?	172 (120 with high baseline CRP)	344 (172 pre, 172 post)	Videos	Yes	Yes	Criterion validity; construct validity; responsiveness
Green 2011	SES-CD	?	18 (15 with high HBI)	18	Chart re-view of colonoscopy	?	Yes	Criterion validity; construct validity
Hebuterne 2013	CDEIS	4 central readers, 1 local	89 (44 with paired endoscopy)	88 (44 pre, 44 post) 2 central reviews per endoscopy 133 videos with consent (49 at baseline, 51 at week 10, 33 at week 54) Each video was read by a pair of central readers (266 reads)	Videos	Yes	Yes	Reliability; responsiveness
Jalocha 2009	CDEIS	?	20 patients (10 infliximab, 10 adalimumab)	40 scopes (20 pre, 20 post)	Colonoscopy	N/A - single review	?	Responsiveness
Jones 2008	SES-CD	?	164 patients	157 with SES-CD scored	Colonoscopy	?	?	Criterion validity; construct validity
Khanna 2016	SES-CD CDEIS	4 central readers	50 patients	4 central readers performed 3 independent measurements for each of the 50 videos Each of the 50 videos was replicated twice to create 150 samples that	Videos	Yes	Yes	Reliability; construct validity

Table 1. Summary of Included Studies (Continued)

				were read by the four readers in random order				
Lobaton Ortega 2013	SES-CD	≥ 2	115 patients (SES-CD)	115 colonoscopies performed; 89 were complete ileocolonoscopies	Colonoscopy	?	Yes	Criterion validity; construct validity
	CDEIS		46 patients (CDEIS)	46 colonoscopies performed				
Mary 1989	CDEIS	≤ 29	a) 75	a) Reproducibility=75 patients (75 colonoscopies read twice = 150 reviews)	Colonoscopy-one reader performing the	Yes	?	Reliability; construct validity; responsiveness
			b) 54	b) Responsiveness=54 patients (103 colonoscopies read once)	procedure, one view-ing through teaching tube			
Molander 2013	SES-CD	≥ 8 (local)	183 (109 Infliximab, 74 adalimumab)	183	Colonoscopy	N/A - single review	?	Criterion validity
Reinisch 2010	SES-CD	?	62	186 with Baseline, week 12 and week 24 results	?	?	?	Criterion validity
Rimola 2011	CDEIS	2	48	48 (single review but both reviewers involved: one performed examination and other annotated findings)	Colonoscopy	N/A - single review	Yes	Criterion validity
Rutgeerts 2002	CDEIS	?	66 patients (24 placebo, 19 infliximab 5mg/kg, 23 infliximab 10 mg/kg)	198 (baseline, Wk 10, Wk 54 for 66 patients)	Colonoscopy	N/A - single review	?	Construct validity; responsiveness

Table 1. Summary of Included Studies (Continued)

Rutgeerts 2006	CDEIS	> 1	99 (75 evaluated)	231 (99 baseline, 74 week 10, 58 week 54)	Colonoscopy - by local endoscopist, video - by single reader for mucosal healing	N/A - single review	Yes	Responsiveness
Rutgeerts 2010	SES-CD	?	62	180 (baseline, week 12 and week 52)	Colonoscopy	?	?	Criterion validity;
Rutgeerts 2016	SES-CD CDEIS	1 blind-ed central reader up to 2 designated endoscopists per site	Baseline: 129 Week 12: 122 Week 52: 84	Baseline: 129 Week 12: 122 Week 52: 84	Colonoscopy and videos	Yes	Yes	Reliability;
Schoepfer 2010	SES-CD	5	122 (plus 43 healthy controls)	140 colonoscopies	Colonoscopy	N/A - single review	Yes	Criterion validity; construct validity
Sipponen 2008a	CDEIS	4	77	106 procedures (25 patients with 2 scopes, 2 patients with 3)	Colonoscopy	N/A - single review	?	Criterion validity; construct validity
Sipponen 2008b	SES-CD	?	61	87 ileocolonoscopy in 61 patients (37 scoped once, 22 twice, 2 three times) Single read of each	Colonoscopy	N/A - single review	? (blind to fecal assay results but unclear if they were aware of CDAI score)	Criterion validity; construct validity
Sipponen 2008c	CDEIS	4	15 (14 infliximab, 1 adalimumab)	30 colonoscopies (baseline and week 12, or week 10 for adalimumab patients)	Colonoscopy	N/A - single review	?	Criterion validity; construct validity; responsiveness

Table 1. Summary of Included Studies *(Continued)*

Sipponen 2010a	SES-CD	3	19	38 (19 at baseline and 19 at 4-6 months)	Colonoscopy	N/A - single review	?	Criterion validity; responsiveness
Sipponen 2010b	SES-CD	4	86, 32 with second endoscopy	86 baseline, 32 follow-up	Colonoscopy	N/A - single review	?	Criterion validity; construct validity; responsiveness
Vieira 2009	CDEIS	2	38	38 double balloon enteroscopies 2 endoscopists reviewed each by video	Video	N/A - single review	Yes	Criterion validity

Table 2. Scoring Index Cut-offs

Study ID	Index	Outcome	Cut-off	Index	Outcome	Cut-off
Af Bor-jorkesten 2012	SES-CD	Remission	Remission 0-2; Mild 3-6; Moderate 7-15; Severe > 16			
Bamba 2012	SES-CD	Remission	Musocal healing 0-2			
Ferrante 2010	SES-CD	Response	Decrease of at least 50% and at least 5 points	CDEIS	Response	Decrease of at least 50% and at least 5 points
Ferrante 2013	SES-CD	Response	Decrease of at least 50% and at least 5 points	CDEIS	Response	Decrease of at least 50% and at least 5 points
Hebuterne 2013	CDEIS	Remission	Remission: < 6 Complete Remission: < 3	CDEIS	Response	CDEIS de- crease of > 5 points
Jones 2008	SES-CD	Remission	Cutoff > 7, predicted by biomarkers			
Molander 2013	SES-CD	Remission	Remission 0-2; Mild 3-6; Moderate 7-15; Severe $\geq$ 16			
Reinisch 2010	SES-CD	Remission	Cutoff > 5 for predicting future remission			
Schoepfer 2010	SES-CD	Remission	Inactive:0-3, Mild 4-10, Moderate 11-19, High $\geq$ 20			
Sipponen 2008a	CDEIS	Remission	Inactive:0-2, Mild 3-9, Moderate 9-12, Severe $\geq$ 12			
Sipponen 2008b	SES-CD	Remission	Suggested cutoff 3 for inactive,			

Table 2. Scoring Index Cut-offs *(Continued)*

but still unsettled						
Sipponen 2008c	CDEIS	Remission	Inactive:0-2, Mild 3-9, Moderate 9-12, Severe ≥ 12	CDEIS	Response	1-class improvement
Sipponen 2010a	SES-CD	Remission	Inactive:0-2, Mild 3-6, Moderate 7-15, Severe ≥ 16	SES-CD	Response	Response: ≤ 2 or 2-3 class decrease Partial response: class improvement
Sipponen 2010b	SES-CD	Remission	Inactive:0-2, Mild 3-6, Moderate 7-15, Severe ≥ 16			
Vieira 2009	CDEIS	Remission	CDEIS ≥ 3 considered positive intestinal inflammation			

Table 3. Reliability

Study ID	Index	Inter-rater kappa (between raters)	Inter-rater ICC (between raters)	Intra-rater kappa (within rater)	Intra-rater ICC (within rater)	Internal Consistency
Daperno 2004	SES-CD	Ulcers: Ileum: 0.884, Right Colon: 0.919, Transverse Colon: 0.972, Left Colon: 0.857, Rectum: 0.819 Ulcerated Surface: Ileum: 0.882, Right Colon: 0.811, Transverse Colon: 0.942, Left Colon: 0.878, Rectum: 0.857	ICC: 0.9815 (95% CI 0.9705 to 0.9884)			

Table 3. Reliability (Continued)

		<p>Affected Surface:</p> <p>Ileum: 0.811,</p> <p>Right Colon: 0.791,</p> <p>Transverse Colon: 0.841,</p> <p>Left Colon: 0.909,</p> <p>Rectum: 0.838</p> <p>Stenosis:</p> <p>Ileum: 1.000,</p> <p>Right Colon: 1.000,</p> <p>Transverse Colon: 1.000,</p> <p>Left Colon: 1.000,</p> <p>Rectum: NA</p>		
Daperno 2014	SES-CD	<p>Average measures: ICC 0.994 (95% CI 0.976-1.000)</p> <p>Single measures: ICC 0.929 (95% CI 0.757-0.98)</p>		
Khanna 2016	SES-CD	<p>0.83 (95% CI 0.75 to 0.88)</p> <p>0.91 (95% CI 0.89 to 0.95)</p>		
Rutgeerts 2016	SES-CD	<p>Baseline: 0.77 (95% CI 0.69 to 0.83)</p> <p>Week 12: 0.86 (95% CI 0.81 to 0.90)</p> <p>Week 52: 0.82 (95% CI 0.73 to 0.88)</p>		
Daperno 2004	CDEIS	<p>Deep Ulcers:</p> <p>Ileum: 0.683,</p> <p>Right Colon: 0.811,</p> <p>Transverse Colon: 1.000,</p> <p>Left Colon: 0.900,</p> <p>Rectum: 0.666</p> <p>Superficial Ulcers:</p> <p>Ileum: 0.701,</p> <p>Right Colon: 0.727, Trans- verse Colon: 0.738,</p>	<p>ICC: 0.9090 (95% CI 0.8580 to 0.9423)</p>	

Table 3. Reliability (Continued)

Left Colon: 0.628,

Rectum: 0.767

Daperno 2014	CDEIS	Average measures: ICC 0.985 (95% CI 0.939-1.000) Single measures: ICC 0.835 (95% CI 0.540-0.995)	
Hebuterne 2013	CDEIS	Week 0 ICC: 0.60 (95%CI 0.39-0.75) Week 10 ICC: 0.74 (95% CI 0.59-0.84) Week 54 ICC: 0.81 (95% CI 0.65-0.90)	
Khanna 2016	CDEIS	0.71 (95% CI 0.63 to 0.76)	0.89 (95% CI 0.86 to 0.93)
Mary 1989	CDEIS	ICC: 0.96 (p < 0.001)	
Rutgeerts 2016	CDEIS	Baseline: 0.78 (95% CI 0.65 to 0.87) Week 12: 0.92 (95% CI 0.86 to 0.95) Week 52: 0.86 (95% CI 0.73 to 0.93)	

No studies assessed intra-rater reliability

Table 4. Content Validity

Study ID	Index	Methods
N/A	CDEIS*	?
N/A	SES-CD*	?

*no studies assessed content validity

Table 5. Criterion Validity

Study ID	Index	Outcome	Correlation
Af Borjorkesten 2012	SES-CD	CRP	r = 0.56 (p < 0.001)
	SES-CD	Calprotectin	r = 0.56 (p < 0.001)
Bamba 2012	SES-CD	CRP	r = 0.621 (p = ?)

Table 5. Criterion Validity (Continued)

	SES-CD	ESR	$r = 0.399$ ($p = ?$)
Daperno 2004	SES-CD	CRP	$r = 0.472$ ($p < 0.001$)
	SES-CD	Other	IBDQ, serum albumin, BMI
Denis 2007	CDEIS	CRP	"no correlation"
	CDEIS	Calprotectin	"no correlation"
D'Haens 1999	CDEIS	CRP	CRP change: $r = 0.47$ ($p = 0.11$)
Ferrante 2013	SES-CD	CRP	CRP change: $r = 0.28$ ($p = 0.0002$)
	CDEIS	CRP	CRP change: $r = 0.27$ ($p = 0.0004$)
Green 2011	SES-CD	CRP	$r = 0.68$ ($p = 0.02$)
	SES-CD	ESR	"did not correlate" ($p = 0.6$)
	SES-CD	Other	Ileal SES-CD to Ileal CTE: $r = 0.69$ ($p = 0.007$)
Jones 2008	SES-CD	CRP	$r = 0.46$ ($p < 0.05$)
	SES-CD	Calprotectin	$r = 0.45$ ($p < 0.05$)
	SES-CD	Lactoferrin	$r = 0.48$ ($p < 0.05$)
Lobaton Ortega 2013*	SES-CD	CRP	* $r = 0.650$ ($p < 0.001$)
	SES-CD	Calprotectin	* $r = 0.807$ ($p < 0.001$)
	SES-CD	ESR	* $r = 0.397$ ($p = 0.015$)
	CDEIS	CRP	* $r = 0.650$ ($p < 0.001$)
	CDEIS	Calprotectin	* $r = 0.807$ ($p < 0.001$)
	CDEIS	ESR	* $r = 0.397$ ($p = 0.015$)
Molander 2013	SES-CD	CRP	$r = ?$ SES-CD ≤ 2 and PGA = 0 showed difference in CRP vs. active disease ($p < 0.001$)
	SES-CD	Calprotectin	$r = ?$ SES-CD ≤ 2 and PGA = 0

Table 5. Criterion Validity (Continued)

			showed difference in calprotectin vs. active disease ($p < 0.001$)
Rutgeerts 2010	SES-CD	Other	<p>CDAI change: 1 point increase in CDAI at week 12 associated with 5 point increase in CDAI by week 52 ($p < 0.05$)</p> <p>IBDQ change: 1 point decrease in SES-CD at week 12 associated with 2 point decrease in IBDQ at week 52 ($p < 0.05$)</p>
Rimola 2011	CDEIS	Other	MaRIA score: $r = 0.8$ ($p < 0.001$)
Schoepfer 2010	SES-CD	CRP	$r = 0.53$ ($p < 0.01$)
	SES-CD	Calprotectin	$r = 0.75$ ($p < 0.01$)
	SES-CD	Other	Leukocytes: $r = 0.42$ ($p < 0.01$)
Sipponen 2008a	CDEIS	CRP	$r = 0.553$ ($p < 0.001$)
	CDEIS	Calprotectin	$r = 0.729$ ($p < 0.001$)
	CDEIS	Lactoferrin	$r = 0.773$ ($p < 0.001$)
Sipponen 2008b	SES-CD	Calprotectin	<p>Ileocolonic/colonic: $r = 0.642$ ($p < 0.001$)</p> <p>Ileal/upper GI: $r = 0.317$ (NS)</p>
	SES-CD	Lactoferrin	<p>Ileocolonic/colonic: $r = 0.627$ ($p < 0.001$)</p> <p>Ileal/upper GI: $r = 0.180$ (NS)</p>
Sipponen 2008c	CDEIS	CRP	$r = 0.608$ ($p < 0.001$)
	CDEIS	Calprotectin	$r = 0.831$ ($p < 0.001$)
	CDEIS	Lactoferrin	$r = 0.865$ (0.001)
Sipponen 2010a	SES-CD	Calprotectin	<p>$r = ?$</p> <p>If SES-CD response: median calprotectin decrease from 1282 mg/g (range 156-2277 mg/g) to 73 mg/g (range 7-222; $p = 0.005$)</p> <p>If SES-CD non-response: median calprotectin decrease from 1017 mg/g (range 53-3928 mg/g) to 223 mg/g (range 35-15330 mg/g; $p = 0.594$)</p>
	SES-CD	Lactoferrin	<p>$r = ?$</p> <p>If SES-CD response: median lactoferrin decrease from 223 mg/g (range 2.8-802 mg/g) to 0.6 mg/g (range 0.0-420 mg/g; $p = 0.005$)</p>

Table 5. Criterion Validity (Continued)

If SES-CD non-response: median lactoferrin decrease from 22.5 mg/g (range 2.1-629) to 13.0 mg/g (range 3.5-1259; p = 0.515)			
Sipponen 2010b	SES-CD	Calprotectin	$r = 0.525$ ($p < 0.001$)
	SES-CD	CRP	$r = ?$ ($p > 0.05$)
	SES-CD	Other	CDAI change: $r = 0.282$ ($p = 0.118$)
	CDEIS	CRP	$r = 0.521$ ($p < 0.001$) CRP change: $r = ?$ ($p > 0.05$)
	CDEIS	Other	CDAI change: $r = 0.240$ ($p = 0.185$)
Vieira 2009	CDEIS	Calprotectin	? ($p < 0.001$)
	CDEIS	Lactoferrin	? ($p < 0.001$)

NS = Not significant

*Loboton 2012 unclear if "endoscopic activity" refers to SES-CD or CDEIS

Table 6. Construct Validity

Study ID	Index	Comparison	Correlation
Af Borjorkesten 2012	SES-CD	CDAI	0.4 ($p < 0.001$)
	SES-CD	Harvey Bradshaw Index	$r = 0.32$ ($p < 0.001$)
Bamba 2012	SES-CD	Modified Rut-geerts Score	$r = 0.748$ (p value ?)
Colombel 2003	CDEIS	CDAI	"positively correlated" $r = ?$ ($p = 0.034$)
Daperno 2004	SES-CD	CDAI	0.39 ($p < 0.001$)
	SES-CD	CDEIS	$r = 0.357$ ($p < 0.001$) CDEIS multiple regression correlation coefficient: 0.920 (95% CI 0.8740, 0.9497) Spearman rank correlation coefficients: 0.887 (95% CI 0.842, 0.920) and 0.910 (95% CI 0.873, 0.936)
Denis 2007	CDEIS	CDAI	"no correlation" to CDEIS

Table 6. Construct Validity (Continued)

≥ 6 versus <6			
D'Haens 1999	CDEIS	CDAI	Δ CDAI to Δ CDEIS $r = 0.56$ ($p = 0.002$)
Ferrante 2010	SES-CD	CDEIS	$r = 0.810$ ($p < 0.001$)
Ferrante 2013	SES-CD	CDAI	Δ CDAI to Δ SES-CD $r = 0.27$ ($p = 0.0003$)
	SES-CD	CDEIS	Δ CDEIS to Δ SES-CD $r = 0.89$ ($p < 0.0001$)
	CDEIS	CDAI	Δ CDAI to Δ CDEIS $r = 0.22$ ($p < 0.0045$)
Green 2011	SES-CD	HBI	$r = 0.70$ ($p = 0.009$)
Jones 2008	SES-CD	CDAI	$r = 0.15$ ($p > 0.05$)
Khanna 2016	SES-CD	CDAI	$r = 0.92$ (95% CI 0.88 to 0.95)
	SES-CD	GELS	$r = 0.74$ (95% CI 0.64 to 0.81)
	CDEIS	CDAI	$r = 0.92$ (95% CI 0.88 to 0.95)
	CDEIS	GELS	$r = 0.75$ (95% CI 0.67 to 0.81)
Lobaton Ortega 2013	SES-CD	CDEIS	$r = 0.968$ ($p < 0.001$)
	CDEIS	CDAI	$r = 0.442$ ($p = 0.004$)
Mary 1989	CDEIS	GELS	$r = 0.83$ ($p < 0.001$)
Rutgeerts 2002	CDEIS	CDAI	Δ CDAI to Δ CDEIS a) Week 10 $r = 0.307$ ($p = 0.021$) b) Week 54 $r = 0.498$ ($p < 0.001$)
Rutgeerts 2006	CDEIS	CDAI	Δ CDAI to Δ CDEIS a) Week 10 $r = 0.33$ ($p = 0.002$) b) Week 54 $r = 0.37$ ($p = 0.003$)
Schoepfer 2010	SES-CD	CDAI	$r = 0.38$ ($p < 0.01$)
Sipponen 2008a	CDEIS	CDAI	$r = 0.381$ ($p < 0.001$)
Sipponen 2008b	SES-CD	CDAI	0.346 ($p < 0.001$) (Spearman's rank order)
Sipponen 2008c	CDEIS	CDAI	$r = 0.592$ ($p = 0.001$)

Table 6. Construct Validity (Continued)

Sipponen 2010b	SES-CD	CDAI	$r = 0.473$ ($p < 0.0001$)
	SES-CD	CDEIS	$r = 0.938$ ($p < 0.0001$)
	CDEIS	CDAI	$r = 0.446$ ($p < 0.0001$)

Table 7. Responsiveness

Study ID	Index	Treatment	Responsiveness Measure	Association with Responsiveness Measure
D'Haens 1999	CDEIS	Infliximab	CDEIS change after	Mean decrease:
		a) 5 mg/kg	4 weeks	a) 15.1+/- 6.9 to 6.4+/-5.1 ($p < 0.01$)
		b) 10 mg/kg		b) 10.6+/- 7.8 to 4.3+/-5.4 ($p < 0.01$)
		c) 20 mg/kg		c) 13.3+/- 6.9 to 5.2+/-2.8 ($p < 0.01$) Overall: 13.0+/- 7.1 to 5.3+/-4.4 ($p < 0.001$)
Ferrante 2013	SES-CD	Azathioprine and/or Infliximab		Note: Steroid free remission compared for endoscopic response definition but not specific to one EI
	CDEIS	Azathioprine and/or Infliximab		Note: Steroid free remission compared for endoscopic response definition but not specific to one EI
Hebuterne 2013	CDEIS	Certolizumab	a) CDEIS change	a) mean change: 3.79 ($p < 0.0001$)
			b) Week 10	b) mean change 5.7 (95% CI 4.6-6.8)
			c) Week 54	($p < 0.0001$)
				c) mean change 4.7 (95% CI 3.5 to 7.5) ($p < 0.0001$)
Jalocha 2009	CDEIS	TNF-antagonist	CDEIS change	a) mean CDEIS decrease 4.8 ($p = ?$)
			a) Week 10 if Infliximab	b) mean CDEIS decrease 5.068 ($p = ?$)
			b) Week 12 if Adalimumab	
Mary 1989	CDEIS	Oral Prednisone	CDEIS change to change in GELS after treatment	$r = 0.72$ ($p < 0.001$)
Rutgeerts 2002	CDEIS	Infliximab	CDEIS change Week 10	mean change -4.9 with infliximab vs. -2.0 in placebo group ($p = 0.003$)
Rutgeerts 2006	CDEIS	Infliximab	% CDEIS change	% change:

Table 7. Responsiveness (Continued)

			a) Week 10 b) Week 54	a) 76% with 3 doses vs 32% with single dose ($p < 0.001$) b) 93% in scheduled group vs. 54% in episodic group ($p = 0.026$)
Sipponen 2008c	CDEIS	TNF-antagonist	CDEIS change	mean decrease 8.2 ($p = 0.002$)
Sipponen 2010a	SES-CD	Oral therapy enhancement (budesonide, mesalamine, metronidazole or immunomodulator)	SES-CD at 4-6 months	$r = ?$ SES-CD decreased from 12 to 7 ($p = 0.006$)
Sipponen 2010b	SES-CD	"CD therapy enhancement" 13/32 received anti-TNF	SES-CD change	SES-CD decreased from 15 to 5 ($p < 0.001$)
	CDEIS	"CD therapy enhancement" 13/32 received anti-TNF	CDEIS change	CDEIS decreased from 11.4 to 2.7 ($p < 0.001$)

Table 8. Feasibility

Study ID	Index	Feasibility Scoring
N/A	SES-CD	?
N/A	CDEIS	?

No identified studies reported feasibility scoring or item scaling

Table 9. Summary of Operating Properties of Endoscopic Scoring Indices for Crohn's Disease

Scoring index	Validity			Reliability				Respon- siveness	Feasibili- ty
	Content va- lidity	Criterion valid- ity	Construct va- lidity	In- tra-rater	In- ter-rater	Test- retest	Internal consisten- cy		
CDEIS	?	+	+	+	+	?	?	+	?
SES-CD	?	+	+	+	+	?	?	+	?

+ positive rating

? no information or indeterminate rating

- Negative rating

Table 10. The Methodological Quality of Endoscopic Index Measurement Properties as Described in the Original Development Articles (COSMIN Checklist)

	Study ID	Scoring in- dex	Internal consis- tency	Reliabil- ity	Mea- sure- ment error	Content validity	Struc- tural va- lidity (CSV)	Hypoth- esis testing (CSV)	Cross -cultural validity (CSV)	Criteri- on validity	Respon- siveness	Inter- pretabil- ity
1	Af Borjorkesten 2012	SES-CD	?	?	?	?	?	excellent	?	excellent	?	?
2	Bamba 2012	SES-CD	?	?	?	?	?	?	?	fair	?	?
3	Colombel 2003	CDEIS	?	?	?	?	?	poor	?	?	?	?
4	Daperno 2004	SES-CD	?	good	?	?	?	excellent	?	excellent	?	?
5	Daperno 2014	SES-CD	?	good	?	?	?	?	?	?	?	?
		CDEIS										
6	Denis 2007	CDEIS	?	?	?	?	?	poor	?	poor	?	?
7	D'Haens 1999	CDEIS	?	?	?	?	?	fair	?	fair	fair	?
8	Ferrante 2010	SES-CD	?	?	?	?	?	poor	?	?	?	?

Table 10. The Methodological Quality of Endoscopic Index Measurement Properties as Described in the Original Development Articles (COSMIN Checklist) *(Continued)*

		CDEIS										
9	Ferrante 2013	SES-CD	?	?	?	?	?	fair	?	fair	fair	?
		CDEIS										
10	Green 2011	SES-CD	?	?	?	?	?	poor	?	poor	?	?
11	Hebuterne 2013	CDEIS	?	good	?	?	?	?	?	?	good	?
12	Jalocha 2009	CDEIS	?	?	?	?	?	?	?	?	poor	?
13	Jones 2008	SES-CD	?	?	?	?	?	excellent	?	excellent	?	?
14	Khanna 2016	SES-CD	?	excellent	?	?	?	excellent	?	?	?	?
		CDEIS										
15	Lobaton Ortega 2013	SES-CD	?	?	?	?	?	fair	?	excellent	?	?
		CDEIS										
16	Mary 1989	CDEIS	?	excellent	?	?	?	excellent	?	?	excellent	?
17	Molander 2013	SES-CD	?	?	?	?	?	?	?	poor	?	?
18	Reinisch 2010	SES-CD	?	?	?	?	?	?	?	poor	?	?
19	Rimola 2011	CDEIS	?	?	?	?	?	?	?	fair	?	?
20	Rutgeerts 2002	CDEIS	?	?	?	?	?	good	?	?	good	?
21	Rutgeerts 2006	CDEIS	?	?	?	?	?	?	?	?	excellent	?
22	Rutgeerts 2010	SES-CD	?	?	?	?	?	?	?	good	?	?
23	Rutgeerts 2016	SES-CD	?	?	?	?	?	?	?	excellent	?	?
		CDEIS										
24	Schoepfer 2010	SES-CD	?	?	?	?	?	excellent	?	?	excellent	?
25	Sipponen 2008a	CDEIS	?	?	?	?	?	excellent	?	excellent	?	?

Table 10. The Methodological Quality of Endoscopic Index Measurement Properties as Described in the Original Development Articles (COSMIN Checklist) *(Continued)*

26	Sipponen 2008b	SES-CD	?	?	?	?	?	good	?	good	?	?
27	Sipponen 2008c	CDEIS	?	?	?	?	?	fair	?	fair	fair	?
28	Sipponen 2010a	SES-CD	?	?	?	?	?	?	?	poor	fair	?
29	Sipponen 2010b	SES-CD	?	?	?	?	?	good	?	poor	good	?
		CDEIS										
30	Vieira 2009	CDEIS	?	?	?	?	?	?	?	poor	?	?

CSV = construct validity

? no information or indeterminate rating

APPENDICES

Appendix 1. Search strategies

MEDLINE

1. exp Crohn disease/
2. Crohn* disease.mp.
3. Exp Inflammatory bowel disease/
4. Inflammatory bowel disease.mp.
5. IBD.mp.
6. Or/1-5
7. Crohn* disease endoscopic index of severity.mp.
8. CDEIS.mp.
9. Simple endoscopic scale for Crohn* disease.mp.
10. SES-CD.mp
11. Rutgeerts scor*.mp.
12. (mucosal adj2 healing).mp.
13. (mucosal adj2 improvement).mp.
14. (endoscop* adj2 improv*).mp.
15. (endoscop* adj3 respon*).mp.
16. endoscopic remission.mp.
17. deep remission.mp.
18. Or 7-17
19. 6 and 18

EMBASE

1. exp Crohn disease/
2. Crohn* disease.mp.
3. Exp Inflammatory bowel disease/
4. Inflammatory bowel disease.mp.
5. IBD.mp.
6. Or/1-5
7. Crohn* disease endoscopic index of severity.mp.
8. CDEIS.mp.
9. Simple endoscopic scale for Crohn* disease.mp.
10. SES-CD.mp
11. Rutgeerts scor*.mp.

12. (mucosal adj2 healing).mp.
13. (mucosal adj2 improvement).mp.
14. (endoscop* adj2 improv*).mp.
15. (endoscop* adj3 respon*).mp.
16. endoscopic remission.mp.
17. deep remission.mp.
18. Or 7-17
19. 6 and 18

Cochrane Library

- #1 MeSH descriptor: [Crohn Disease] explode all trees
- #2 MeSH descriptor: [Inflammatory Bowel Diseases] explode all trees
- #3 #1 or 2
- #4 Crohn* disease endoscopic index of severity.mp.
- #5 CDEIS.mp.
- #6 Simple endoscopic scale for Crohn* disease.mp.
- #7 SES-CD.mp
- #8 Rutgeerts scor*.mp.
- #9 mucosal healing
- #10 mucosal improvement
- #11 endoscopic improvement
- #12 endoscopic response
- #13 endoscopic remission
- #14 deep remission
- #15 or/#4-14
- #16 #3 and #15

IBD Specialized Register

Title or abstract: "CDEIS" or "SES" or "mucosal" or "endoscopic" or "Rutgeerts score" or "deep remission"

CONTRIBUTIONS OF AUTHORS

Development of study concept and design: RK, SAN, BGF, GDH, WJS, GYZ, JKM, VJ, BGL

Study supervision: RK, CEP, JKM

Acquisition, analysis, and interpretation of the data: RK, SAN, JKM, CEP

Drafting of the manuscript: RK, SAN, BGF, GDH, WJS, GYZ, JKM, CEP, VJ, BGL

DECLARATIONS OF INTEREST

Sigrid A Nelson, John K MacDonald, Claire E Parker and GY Zou have no known conflicts.

Reena Khanna has received consulting fees from AbbVie, Janssen, Pfizer, Shire, and Takeda.

Endoscopic scoring indices for evaluation of disease activity in Crohn's disease (Review)

Copyright © 2016 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Brian Feagan has served on the board of Abbott/AbbVie, Amgen, Astra Zeneca, Avaxia Biologics Inc., Bristol-Myers Squibb, Celgene, Centocor Inc., Elan/Biogen, Ferring, JnJ/Janssen, Merck, Novartis, Novonordisk, Pfizer, Prometheus Laboratories, Protagonist, Salix Pharma, Takeda, Teva, Tillotts Pharma AG, and UCB Pharma; has received consultancy fees from Abbott/AbbVie, Actogenix, Albireo Pharma, Amgen, Astra Zeneca, Avaxia Biologics Inc., Axcan, Avir Pharma, Baxter Healthcare Corp., Biogen Idec, Boehringer-Ingelheim, Bristol-Myers Squibb, Calypso Biotech, Celgene, Elan/Biogen, EnGene, Ferring Pharma, Roche/Genentech, GiCare Pharma, Gilead, Given Imaging Inc., GSK, Ironwood Pharma, Janssen Biotech (Centocor), JnJ/Janssen, Kyowa Kakko Kirin Co Ltd., Lexicon, Lilly, Merck, Millennium, Nektar, Novonordisk, Pfizer, Prometheus Therapeutics and Diagnostics, Protagonist, Receptos, Salix Pharma, Serono, Shire, Sigmoid Pharma, Synergy Pharma Inc., Takeda, Teva Pharma, Tillotts, UCB Pharma, Vertex Pharma, Warner-Chilcott, Wyeth, Zealand, and Zyngenia; and has received lecture fees from Abbott/AbbVie, JnJ/Janssen, Takeda, Warner-Chilcott, and UCB Pharma.

Geert D'Haens has received consultancy fees from Abbvie, Ablynx, Amakem, AM Pharma, Avaxia, Biogen, Bristol Meiers Squibb, Boehringer Ingelheim, Celgene, Celltrion, Cosmo, Covidien, Ferring, DrFALK Pharma, Engene, Galapagos, Gilead, Glaxo Smith Kline, Hospira, Immunic, Johnson and Johnson, Lycera, Medimetrics, Millenium/Takeda, Mitsubishi Pharma, Merck Sharp Dome, Mundipharma, Novonordisk, Pfizer, Prometheus laboratories/Nestle, Protagonist, Receptos, Robarts Clinical Trials, Salix, Sandoz, Setpoint, Shire, Teva, Tigenix, Tillotts, Topivert, Versant and Vifor; has held grants/grants pending from Abbvie, Covidien, Ferring, DrFALK Pharma, Millenium/Takeda, Merck Sharp Dome, Mundipharma, Pfizer, and Prometheus laboratories/Nestle; has received payment for lectures from Abbvie, Ferring, Johnson and Johnson, Merck Sharp Dome, Mundipharma, Norgine, Pfizer, Shire, Millenium/Takeda, Tillotts and Vifor; and has stock/stock options with Engene.

William J Sandborn has received consultancy fees from Abbott Laboratories, ActoGenix NV, AGI Therapeutics, Inc., Alba Therapeutics Corporation, Albireo,

Alfa Wasserman, Amgen, AM-Pharma BV, Anaphore, Astellas Pharma, Athersys, Inc., Atlantic Healthcare Limited, Axcan Pharma (now Aptalis), BioBalance Corporation, Boehringer-Ingelheim Inc, Bristol Meyers Squibb: (both money paid to WS and institution), Celegene, Celek Pharmaceuticals, Cellerix SL, Cerimon Pharmaceuticals, ChemoCentryx, CoMentis, Cosmo Technologies, Coronado Biosciences, Cytokine Pharmasciences, Eagle Pharmaceuticals,

Eisai Medical Research Inc., Elan Pharmaceuticals: (both money paid to WS and institution), EnGene, Inc., Eli Lilly, Enteromedics: (both money paid to WS and institution), Exagen Diagnostics, Inc., Ferring Pharmaceuticals, Flexion Therapeutics, Inc., Funxional Therapeutics Limited, Genzyme Corporation, Genentech (now Roche): (both money paid to WS and institution), Gilead Sciences, Given Imaging, Glaxo Smith Kline, Human Genome Sciences, Ironwood Pharmaceuticals (previously Microbia Inc.), Janssen (previously Centocor): (both money paid to WS and institution), KaloBios Pharmaceuticals, Inc., Lexicon Pharmaceuticals, Lycera Corporation, Meda Pharmaceuticals (previously Alaven Pharmaceuticals), Merck Research Laboratories, MerckSerono, Millennium Pharmaceuticals (subsequently merged with Takeda): (both money paid to WS and institution), Nisshin Kyorin Pharmaceuticals Co., Ltd., Novo Nordisk A/S,

NPS Pharmaceuticals, Optimer Pharmaceuticals, Orexigen Therapeutics, Inc., PDL Biopharma: (money paid to institution), Pfizer: (both money paid to WS and institution), Procter and Gamble: (both money paid to WS and institution), Prometheus Laboratories, ProtAb Limited, Purgenesis Technologies, Inc., Relypsa, Inc., Salient Pharmaceuticals, Salix Pharmaceuticals, Inc., Santarus, Schering Plough Corporation (acquired by Merck), Shire Pharmaceuticals: (money paid to institution), Sigmoid Pharma Limited, Sirtris Pharmaceuticals, Inc. (a GSK company), S.L.A. Pharma (UK) Limited, Takeda: (both money paid to WS and institution), Targacept, Teva Pharmaceuticals, Therakos, Tillotts Pharma AG (acquired by Zeria Pharmaceutical Co., Ltd), TxCell SA, UCB Pharma: (both money paid to WS and institution), Viamet Pharmaceuticals, Vascular Biogenics Limited (VBL), Warner Chilcott UK Limited, and Wyeth (now Pfizer); fees for expert testimony from Dickinson, Prud'Homme, Adams & Ingram; grants/grants pending from Abbott Laboratories, Bristol Meyers Squibb, Genentech, Glaxo Smith Kline, Janssen (previously Centocor), Millennium Pharmaceuticals (now Takeda), Novartis, Pfizer, Procter and Gamble Pharmaceuticals, Shire Pharmaceuticals, and UCB Pharma; payment for lectures from Abbott Laboratories, Bristol Meyers Squibb and Janssen (previously Centocor); and holds the following patents: Sandborn WJ. Use of topical azathioprine to treat inflammatory bowel disorders. United States patent number: 5,691,343. Date of patent: November 25, 1997., Sandborn WJ, Rhodes J. Colonic delivery of nicotine to treat inflammatory bowel disease. South African patent number: 97/1020. Date of patent: January 28, 1998., Sandborn WJ. Use of azathioprine to treat Crohn's disease. United States patent number: 5,733,915. Date of patent: March 31, 1998., Sandborn WJ, Rhodes J. Colonic delivery of nicotine to treat inflammatory bowel disease. United States patent number: 5,846,983. Date of patent: December 8, 1998., Sandborn WJ. Azathioprine compositions for colonic administration. New Zealand patent number: 306062. Date of Patent: February 11, 1999., Sandborn WJ. Azathioprine compositions for colonic administration. Singapore patent number: 45647. Date of Patent: March 14, 1999., Sandborn WJ, Rhodes J, Rhodes P, Evans BK. Colonic delivery of nicotine to treat inflammatory bowel disease. United States patent number: 5,889,028. Date of patent: March 30, 1999., Sandborn WJ. Topical formulations of azathioprine to treat inflammatory bowel disorders. United States patent number: 5,905,081. Date of Patent: May 18, 1999., Sandborn WJ. Azathioprine compositions for colonic administration. Australia patent number: 707168. Date of Patent: October 14, 1999. Sandborn WJ, Rhodes J, Evans BK. Intestinal absorption of nicotine to treat nicotine responsive conditions. Australia patent number: 718052. Date of patent: July 20, 2000., Sandborn WJ, Rhodes J. Colonic delivery of nicotine to treat inflammatory bowel disease. United States patent number: 6,166,044. Date of patent: December 26, 2000., Sandborn WJ. Use of topical azathioprine and thioguanine to treat colorectal adenomas. United States patent number: 6,166,024. Date of patent: December 26, 2000., Rhodes J, Evans BK, Rhodes P, Sandborn WJ. Intestinal absorption of nicotine to treat nicotine responsive conditions. United States patent number: 6,238,689. Date of patent: May 29, 2001., Sandborn, WJ. Azathioprine compositions for colonic administration. Czech Republic patent number: 290428. Date of patent: May 27, 2002., Sandborn, WJ, Rhodes J. Colonic delivery of nicotine to treat IBD. Mexico patent number: 209636. Date of Patent August 12, 2002., Sandborn WJ. Enema and enterically-coated oral dosage forms of azathioprine. United States Patent No.: 6,432,967. Date of patent: August 13, 2002., Sandborn WJ, Rhodes J. Colonic delivery of nicotine to treat nicotine responsive conditions. Europe patent number: 0954337. Date of patent: November 2, 2002., Sandborn

WJ, Rhodes J, Rhodes P, Evans BK. Colonic delivery of nicotine to treat IBD. Europe patent number: 893998. Date of patent: April 15, 2003., Sandborn WJ, Rhodes J, Rhodes P, Evans BK. Colonic delivery of nicotine to treat inflammatory bowel disease. Hong Kong patent number: HK1019043. Date of patent: August 1, 2003., Sandborn WJ, Rhodes J, Rhodes P, Evans BK. Colonic delivery of nicotine to treat IBD. China patent number: ZL97192177. Date of patent: November 12, 2003., Sandborn W, Rhodes J, Rhodes P, Evans B. Colonic delivery of nicotine to treat inflammatory bowel disease. Czech patent number: 293616. Patent date: 2004., Rhodes J, Sandborn WJ, Rhodes P, Evans BK. Colonic deliver of nicotine to treat inflammatory bowel disease. Canada patent number: 2,246,235. Patent date: 2007., Sachetto JP, Sandborn WJ, Tremaine WJ. Pharmaceutical composition for the treatment of inflammatory bowel disease. United States patent number: 7341741. Patent date 2008., Rhodes J, Evans BK, Rhodes P, Sandborn WJ. Intestinal absorption of nicotine to treat nicotine responsive conditions. Canadian patent number: 2,260,909. Patent date 2008., and Levy MJ, Camilleri ML, Murray JA, Sandborn WJ. Obesity treatment and device. United States patent number: 7,803,195 B2. Date of patent September 28, 2010.

Vipul Jairath has received scientific advisory board fees from Abbvie, Sandoz, Ferring and Janssen; lecture fees from Takeda and Ferring; and travel support for conference attendance from Vifor pharmaceuticals.

Barrett Levesque has received fees for consultancy from Santarus, Prometheus Labs, and Castlight Health; lecture fees from Warner Chilcott, Abbott Labs, Salix and UCB Pharma; and travel support from Curatio CME, and the American Academy of Insurance Medicine.

All of the aforementioned financial activities are outside the submitted work.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

The methods for assessing the risk of bias in included studies were modified from the protocol. Originally, we planned on using four items to assess risk of bias: blinded design, independent observation, performance bias and detection bias. However, since this is a review of scoring indices rather than interventions, the last two items are not applicable. We chose to assess blinded design and independent observation, and to further assess risk of bias using a system based on the COSMIN tool.

INDEX TERMS

Medical Subject Headings (MeSH)

*Endoscopy, Gastrointestinal; *Severity of Illness Index; Blood Sedimentation; C-Reactive Protein [analysis]; Cohort Studies; Crohn Disease [*pathology]; Feces [chemistry]; Randomized Controlled Trials as Topic; Reproducibility of Results

MeSH check words

Humans