

ORIGINAL ARTICLE

Appraisal of current surgical guidelines for inflammatory bowel disease using the AGREE-S instrument: A scoping review

Zarnigar Mussarat Khan^{1,2}  | Camille Ball¹ | Dalha Saeed¹ | Grace Tai¹ |
Shaneil Chandran¹ | Abhishek Vashista¹ | Simon Davey² | Matthew James Lee³ |
Steven R. Brown^{1,2}  | Daniel Hind⁴ | Adele Elizabeth Sayers^{1,2}

¹The Medical School, University of Sheffield, Sheffield, UK

²Academic Directorate of General Surgery, Sheffield Teaching Hospitals, Sheffield, UK

³Institute of Applied Health Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

⁴School of Medicine and Population Health, University of Sheffield, Sheffield, UK

Correspondence

Zarnigar Mussarat Khan, Academic Directorate of General Surgery, Sheffield Teaching Hospitals, Herries Road, Sheffield S5 7AU, UK.
Email: zarni_k@hotmail.com;

Abstract

Aim: Guidelines play a crucial role in improving patient care by providing clinicians with up to date evidence-based recommendations. A vast number of guidelines exist on the surgical management of inflammatory bowel disease (IBD). The aim of this scoping review was to identify current surgical IBD guidelines, assess their quality and identify areas of variation between the existing guidelines.

Method: A systematic search of the literature from January 2008 to September 2023 was conducted. After identifying eligible guidelines, they were assessed for quality using the Appraisal of Guidelines for Research and Evaluation for Surgical Interventions (AGREE-S) instrument. Data were extracted on descriptive guideline characteristics and recommendations.

Results: Fifteen guidelines were identified globally. Most guidelines were published between 2011 and 2023, with six focusing solely on Crohn's disease, five on ulcerative colitis and four on both. Six guidelines focused exclusively on surgical management, while nine contained both medical and surgical recommendations. The overall mean AGREE-S score was 59%, with more recent guidelines scoring higher.

Conclusions: The quality of IBD surgical guidelines varies considerably. High-quality, collaborative, international guidelines are needed to reduce duplication and ensure consistent, evidence-based surgical care for IBD patients worldwide. Future guideline development should adhere to the AGREE-S criteria to enhance methodological rigour and transparency.

KEYWORDS

AGREE-S, guidelines, IBD

INTRODUCTION

In recent years there has been a notable surge in research activity focused on inflammatory bowel disease (IBD). Comprising ulcerative colitis (UC) and Crohn's disease (CD), nearly 5 million cases were

reported worldwide in 2019 [1]. Both occur as a result of chronic inflammation of the gastrointestinal tract, have a genetic predisposition that increases the risk of developing the disease and have no known cure, resulting in substantial morbidity and mortality [2]. Management of IBD includes both medical and surgical modalities.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Colorectal Disease* published by John Wiley & Sons Ltd on behalf of Association of Coloproctology of Great Britain and Ireland.

As can be imagined for diseases of such prevalence, the abundance of literature is enormous with 91 648 publications generated when searching the term 'inflammatory bowel disease' in PubMed (accessed 22 September 2023), highlighting the need for guidelines to disseminate high-quality, reliable evidence.

In the rapidly advancing field of health research, novel and innovative treatments are constantly emerging requiring refinement of existing practice. Guidelines exist as a set of recommendations designed to aid decision-making and best advice on optimal patient care. Guidelines therefore play a key role in improving patient care and health outcomes by allowing clinicians to make evidence-based decisions in a timely manner [3]. In addition to improving quality of care, guidelines can influence health policy, lead to the development of disease performance measures, promote health equality and improve consistency of care and patient empowerment [4,5].

Recommendations made by guidelines should be informed by systematic review to assess the quality of the research evidence [6]. As well as review, guidelines require transparent development, involvement of key stakeholders and consideration of the benefits and harms of alternative treatments [7]. Guidelines require continual update to take account of new research findings, with reports in the literature that one in five guidelines become out of date 3 years after being published [8]. Furthermore, inconsistency in the reliability and trustworthiness of guidelines has been reported, with over 50% of guidelines not complying with current Institute of Medicine standards [9]. Tools exist to evaluate this variability in the quality of guidelines, such as the AGREE-II (Appraisal of Guidelines, Research and Evaluation II) reporting checklist [10]. More specifically, the Appraisal of Guidelines for Research and Evaluation for Surgical Interventions (AGREE-S) reporting checklist was created to improve the comprehensiveness, completeness and transparency of reporting in surgical guidelines [11].

Guidelines for the medical treatment of IBD were previously appraised for their quality and robustness in 2022 [12]; however, to date there have been no attempts to conduct an appraisal of surgical guidelines. Therefore, the aim of this review was to identify current surgical guidelines for the treatment of IBD, critically appraise the quality of these guidelines using the AGREE-S instrument and identify areas of variation between the existing guidelines.

METHOD

A protocol for this scoping review was developed a priori but has not been published or registered.

Eligibility criteria

Eligibility criteria included articles published between January 2008 and September 2023 that were guidelines as defined by the Institute of Medicine, focusing on the surgical management of IBD (CD and UC) in adults [7]. Only the most up to date version of the guideline

What does this paper add to the literature?

This scoping review highlights that the abundance of literature existing on the surgical management of inflammatory bowel disease in the form of guidelines can vary in quality.

was included. We excluded guidelines that were not in English and those that did not solely focus on IBD but mentioned IBD as a precursor to another condition. Similarly, guidelines focusing on a specific aspect of the disease, for example perianal CD, were excluded. Expert reviews, consensus statements and quality indicators were also excluded. This ensured that the guidelines assessed were comparable in scope, allowing the authors to make better comparisons between guidelines.

Search strategy

We conducted a search of the electronic databases EMBASE and MEDLINE. The search strategy included keywords and terms related to IBD, clinical standards, quality indicators and guidelines tailored to each database. The exact syntax searches are given in Supplementary Material S1. Additionally, we conducted a search of guideline networks and repositories, including Guidelines International Network, CPG InfoBase, the Scottish Intercollegiate Guideline Network and Google Scholar, of which the first 80 pages were searched [13]. No additional sources beyond the electronic databases were searched.

Study selection process

After identifying potentially eligible guidelines, two reviewers (AES and ZK) independently screened the full texts for inclusion. Disagreements were resolved through discussion with a third reviewer (SB). The selection process was guided by the eligibility criteria outlined above.

Data extraction

Descriptive data were extracted by two reviewers (ZK and AES) and included information on organization, country, clinical area (CD or UC) and recommendations.

Guideline appraisal

The quality of the eligible guidelines was assessed using the AGREE-S instrument, which appraises the quality of guidance in the domains of scope and purpose, stakeholders, evidence synthesis, development

of recommendations, editorial independence and implementation and update [14]. Scope and purpose explored whether the guideline was developed according to a protocol, with specific objectives and health questions. The stakeholder domain scores the guideline on its development and target audience. Evidence synthesis reviews the search, selection and quality of evidence included. Development of recommendations scores the guideline on the formulation and implication of the recommendations. Editorial independence examines any potential conflicts of interest. Lastly, the domain of implementation and update considers the barriers to application of the recommendations, as well as procedures for updating and auditing.

Five reviewers (AES, ZK, CB, DS, GT) independently charted data from the included guidelines using a standardized charting form. The form was pilot tested on a sample of included guidelines and revised accordingly. Any disagreements in data charting were resolved through discussion. In cases of significant scoring discrepancies (defined as individual scores differing by more than two points on the seven-point scale for any AGREE-S item), the reviewers discussed their rationale and re-examined the relevant guideline sections together to reach consensus. The final score was determined by averaging the individual scores after discrepancies were resolved. The data charting form included items on guideline characteristics (year of publication, country of publication, focus on UC/CD, surgical/mixed/medical management) and specific recommendations made on surgical interventions.

Five reviewers (AES, ZK, SB, MJL, SD) independently applied the AGREE-S instrument to each guideline to assess quality. Following

submission of scores using the data charting form, the reviewers met to discuss the results where there was significant discrepancy in scores and presented information that may have been overlooked by others.

The scores were scaled to a percentage of the maximum score by domain (maximum of seven points per domain). Recommendation of use was determined as a median of the reviewer's scores for use, whereby 'Yes', 'Yes—with modifications' and 'No' equated to a score of 3, 2 and 1 respectively.

Synthesis of results

The data were not synthesized using a framework, as recommendations were not compatible with the structure of a framework.

RESULTS

Study selection

The searches of bibliographic databases generated 1673 articles. Once duplicates were removed, 1492 articles were screened by title and abstract for inclusion, of which 35 full-text articles were retrieved. Ten of these articles were deemed satisfactory for inclusion. Screening of grey literature resulted in 20 articles being retrieved, of which five guidelines were included. [Figure 1](#) shows a PRISMA diagram for study inclusion.

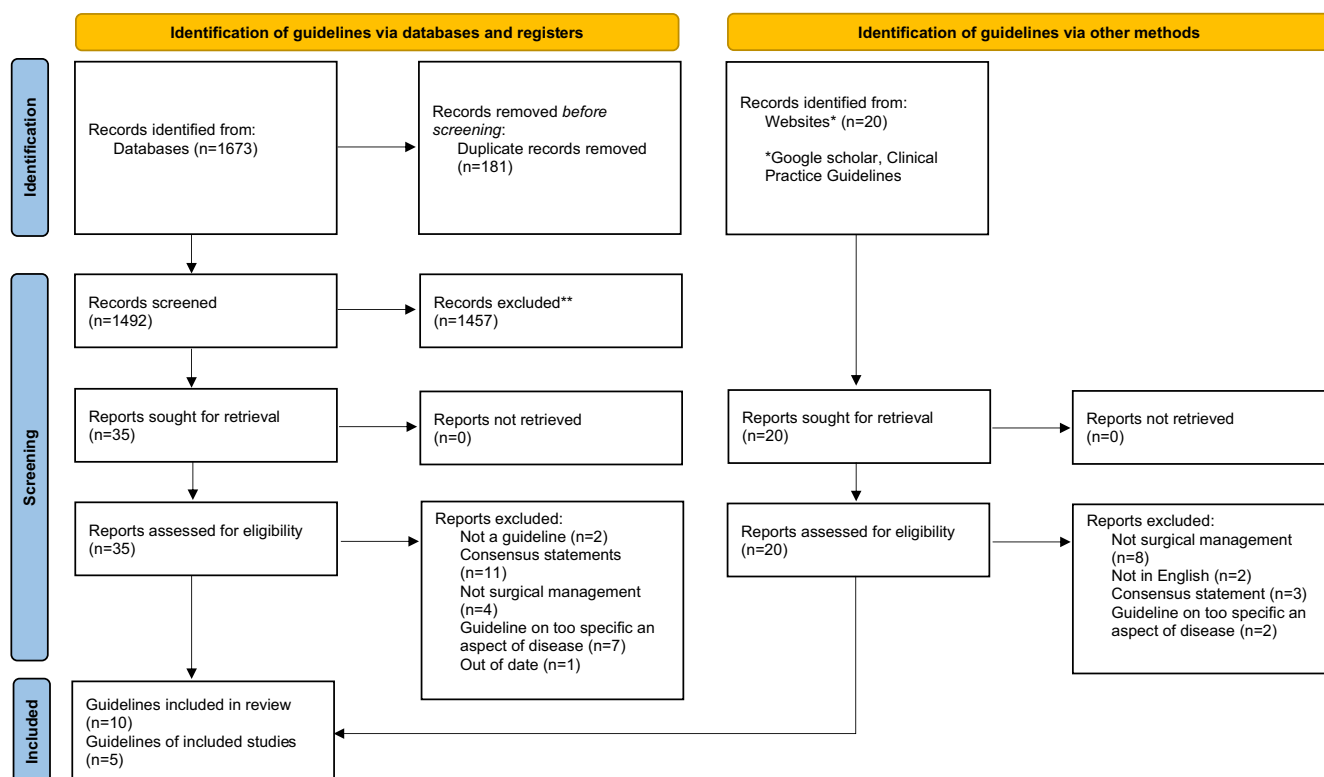


FIGURE 1 PRISMA flow diagram for study inclusion.

Study characteristics

Table 1 displays characteristics of the guidelines included. The included guidelines were published between 2011 and 2023. Of the 15 guidelines, six focused solely on CD, five focused solely on UC and the remaining four focused on the management of IBD including both UC and CD. A total of six guidelines focused solely on surgical management. The remaining nine guidelines contained aspects of both medical and surgical management.

AGREE-S appraisal

Table 2 displays a summary of the characteristics being assessed in the included guidelines based upon the AGREE-S instrument. Of the guidelines reviewed, two were recommended for use, nine were recommended for use with modifications and four were not recommended for use. The mean overall score of all the guidelines reviewed was 59% ($n=15$). For the nine guidelines recommended for use with modifications, the most common reasons were insufficient details on guideline development methodology, lack of clarity on how recommendations were formulated and limited information on implementation and updating procedures. Modifications in these areas would improve adherence to AGREE-S criteria.

The domain where papers scored the lowest when averaged out across all the guidelines reviewed was 'Implementation and update' (45%). This was attributed to an absence of a description of facilitators and barriers to guideline applications. The low scores in the implementation and update domain reflect several common issues identified across the appraised guidelines. Firstly the lack of specific tools or resources provided to aid implementation (e.g. summary documents, algorithms, online resources), insufficient guidance on monitoring/auditing criteria to assess guideline implementation and impact, and lastly vague or missing information on update methodology and timelines. For example, the Eder et al. [18] and Lodyga et al. [23] guidelines, which received some of the lowest scores in this domain, did not provide any information on implementation support, auditing plans or future updating procedures.

The highest scoring domain when averaged out across all reviewed guidelines was 'Editorial independence' (70%), which was attributed to the fact that most guidelines would have undergone external peer review before publication. Higher-scoring articles appraised their evidence with the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) tool or the Oxford Centre for Evidence-Based Medicine (CEBM) tool, which was clearly presented within the guideline [30,31]. They also presented clear methods of evidence searches and recommendation formulation using the Population, Intervention, Comparison and Outcome (PICO) format [32]. The highest scoring guidelines were the European Crohn's and Colitis Organisation (ECCO) guidelines on the surgical treatment of UC (76%) and the UC guidelines produced by the German Society for Metabolic Diseases (76%) [21,27].

Papers published from 2020 onwards tended to score higher, with a mean overall score of 64% ($n=6$) versus those published prior to 2020 with a mean overall score of 52% ($n=9$).

To examine variability in individual appraisals, we calculated the standard deviation of scores for each AGREE-S item within each guideline. The items with the highest average variability across guidelines were item 5 in domain 2 (mean SD 1.6, range 0.8–2.2) and item 3 in domain 3 (mean SD 1.5, range 0.5–2.1), highlighting potential challenges in assessing stakeholder involvement and evidence search/selection methods. Conversely, item 1 in domain 1 (mean SD 0.5, range 0–1.1) and item 3 in domain 6 (mean SD 0.6, range 0–1.2) had the lowest variability, suggesting more consistent evaluation of guideline scope and updating procedures.

Surgical management—UC versus CD

Of the 15 guidelines reviewed, six focused solely on the management of CD, five focused solely on UC management and four made recommendations on both. Guidelines on surgical management alone of IBD ($n=6$) scored more highly, with a mean score of 66%, compared with those making mixed recommendations on medical and surgical management, which had a mean score of 55% ($n=9$), although the difference in scores was not significant. There was negligible difference in mean scores by the region in which the guideline was developed. There was also negligible difference in mean scores for guidelines specifically making recommendations on CD ($n=6$, 61%), UC ($n=5$, 60%) and mixed IBD studies ($n=4$, 55%).

DISCUSSION

Over the 13-year period, a total of 15 unique guidelines on IBD surgery were published in English and included in this appraisal [15–29]. Of these 15, only two are recommended for use without modifications, following appraisal with the AGREE-S appraisal tool [15,21]. Four guidelines were deemed to be not recommended for use following appraisal, while the other nine guidelines were felt to be acceptable for use but only after modification to improve quality. Despite the two highest rated guidelines both scoring 76%, the ECCO guideline on UC was only recommended for use with modifications, whereas the guideline by Kucharzik et al. [21] was recommended for use with no modifications. We hypothesize that this is due to the perceived distribution of scores across domains, with Kucharzik et al. scoring from 55% to 88%, compared with Spinelli et al. scoring less consistently with a range from 44% to 82%.

When analysing the mean scores by region, there was no significant difference in quality. It was noted, however, that those guidelines published after 2020 were of a higher quality than those published before 2020. In recent years, a steady improvement in guideline quality has been observed, attributed to the use of appraisal tools such as AGREE-S [33]. The AGREE-S tool assists journal editors in ensuring that published guidelines meet the high quality

TABLE 1 Guideline characteristics.

Year of publication	Author	Title	Organization	Publication	Country	Disease	Surgical or mixed medical and surgical management
2020	Adamina et al. [15]	ECCO guidelines on therapeutics in Crohn's disease: Surgical treatment	European Crohn's and Colitis Organisation	<i>Journal of Crohn's and Colitis</i>	Europe	CD	Surgical
2015	Bernstein et al. [16]	World Gastroenterology Organisation global guidelines inflammatory bowel disease: Update August 2015	World Gastroenterology Organisation	<i>Journal of Clinical Gastroenterology</i>	Global	IBD	Mixed
2018	Brown et al. [17]	The Association of Coloproctology of Great Britain and Ireland consensus guidelines in surgery for inflammatory bowel disease	Association of Coloproctology of Great Britain and Ireland	<i>Colorectal Disease</i>	UK	IBD	Surgical
2023	Eder et al. [18]	Guidelines for the management of ulcerative colitis	Polish Society of Gastroenterology	<i>Przegląd Gastroenterologiczny</i>	Poland	UC	Mixed
2015	Eliadou et al. [19]	New Zealand Society of Gastroenterology guidelines for the management of refractory ulcerative colitis	New Zealand Society of Gastroenterology	<i>New Zealand Medical Journal</i>	New Zealand	UC	Mixed
2021	Holubar et al. [20]	The American Society of Colon and Rectal Surgeons clinical practice guidelines for the surgical management of ulcerative colitis	American Society of Colon and Rectal Surgeons	<i>Diseases of the Colon and Rectum</i>	USA	UC	Surgical
2019	Kucharzik et al. [21]	Updated S3 guideline ulcerative colitis	German Society for Digestive and Metabolic Diseases	<i>Aktualisierte S3-Leitlinie Colitis Ulcerosa der Deutschen Gesellschaft für Gastroenterologie, Verdauungs- und Stoffwechselkrankheiten</i>	Germany	UC	Mixed
2020	Lightner et al. [22]	The American Society of Colon and Rectal Surgeons clinical practice guidelines for the surgical management of Crohn's disease	American Society of Colon and Rectal Surgeons	<i>Diseases of the Colon and Rectum</i>	USA	CD	Surgical
2021	Lodyga et al. [23]	Guidelines for the management of patients with Crohn's disease. Recommendations of the Polish Society of Gastroenterology and the Polish National Consultant in Gastroenterology	Polish Society of Gastroenterology	<i>Przegląd Gastroenterologiczny</i>	Poland	CD	Mixed
2018	Matsuoka et al. [24]	Evidence-based clinical practice guidelines for inflammatory bowel disease	Japanese Society of Gastroenterology	<i>Journal of Gastroenterology</i>	Japan	IBD	Mixed
2011	Mowat et al. [25]	Guidelines for the management of inflammatory bowel disease in adults	British Society of Gastroenterology	<i>Gut</i>	UK	IBD	Mixed
2017	Park et al. [26]	Second Korean guidelines for the management of Crohn's disease	Korean Association for the Study of Intestinal Diseases	<i>Intestinal Research</i>	South Korea	CD	Mixed
2022	Spinelli et al. [27]	ECCO guidelines on therapeutics in ulcerative colitis: surgical treatment	European Crohn's and Colitis Organisation	<i>Journal of Crohn's and Colitis</i>	Europe	UC	Surgical

(Continues)



TABLE 1 (Continued)

Year of publication	Author	Title	Organization	Publication	Country	Disease	Surgical or mixed medical and surgical management
2015	Strong et al. [28]	Clinical practice guideline for the surgical management of Crohn's disease	American Society of Colon and Rectal Surgeons	<i>Diseases of the Colon and Rectum</i>	America	CD	Surgical
2013	Ueno et al. [29]	Evidence-based clinical practice guidelines for Crohn's disease, integrated with formal consensus of experts in Japan	Japanese Society of Gastroenterology	<i>Journal of Gastroenterology</i>	Japan	CD	Mixed

Abbreviations: CD, Crohn's disease; ECCO, European Crohn's and Colitis Organisation; IBD, inflammatory bowel disease; UC, ulcerative colitis.

TABLE 2 Summary of assessed characteristics.

Year of publication	Author	Scope and purpose	Stakeholder involvement	Evidence synthesis	Development of recommendations	Editorial independence	Implementation and update	Overall score	Recommended for use
2020	Adamina et al. [15]	74	73	93	72	86	46	73	Yes
2015	Bernstein et al. [16]	14	15	14	33	37	28	18	No
2018	Brown et al. [17]	69	71	51	77	64	44	67	Yes—with modifications
2023	Eder et al. [18]	76	37	57	58	68	33	53	Yes—with modifications
2015	Eliadou et al. [19]	35	19	18	28	61	29	29	No
2021	Holubar et al. [20]	64	63	84	61	87	51	65	Yes—with modifications
2019	Kucharzik et al. [21]	70	88	62	76	85	57	76	Yes
2020	Lightner et al. [22]	58	64	84	64	90	47	65	Yes—with modifications
2021	Lodyga et al. [23]	67	41	55	61	51	33	49	No
2018	Matsuoka et al. [24]	65	63	61	57	69	51	63	Yes—with modifications
2011	Mowat et al. [25]	64	88	52	62	77	71	73	Yes—with modifications
2017	Park et al. [26]	68	69	76	60	64	46	59	Yes—with modifications
2022	Spinelli et al. [27]	82	64	76	73	77	44	76	Yes—with modifications
2015	Strong et al. [28]	56	55	65	60	44	36	49	No
2013	Ueno et al. [29]	73	80	71	65	97	57	71	Yes—with modifications
Average per domain		62	59	61	60	70	45	59	

expected of such highly citable work, which plays an integral role in supporting best medical practice and guiding current healthcare policy [34]. It is perhaps the implementation of this tool that has led to a decrease in the publication of poor-quality guidelines in recent years, as noted within this review.

Our findings suggest that IBD guidelines focused solely on surgical management tended to have higher quality scores than those combining medical and surgical aspects, although the difference was not significant. This suggests that authors should consider a move away from large, all-encompassing guidelines towards a more focused guideline, completed to a high standard to help with improved usability. More targeted research is needed to determine if narrower scope consistently leads to higher quality and utility. Factors such as guideline length, depth of topic coverage and end-user preferences should be considered to inform optimal guideline design.

The guidelines were developed within multiple regions around the world, with three being international and the remaining 12 being developed within single countries. Despite good-quality guidelines on the surgical management of CD being produced from the ECCO in 2020, a single-country guideline of poor quality was produced the following year within Europe on the same topic. This was replicated in the 2022 ECCO guidelines on UC, again with the production of a low-quality guideline on the same topic by the same European country. This does raise the question: should there be better way of registering guidelines to reduce unnecessary replication, for example with the Guideline International Network [35]? Or, is the issue that some countries do not believe international guidelines are representative of their patient populations or take into account differing healthcare economies? Guideline developers should take into account facilitators and barriers to a guideline's implementation to account for this, unfortunately this is not always the case.

The poorest scoring domain for all guidelines was 'Implementation and update.' While providing a timeline for planned guideline updates is straightforward, identifying potential facilitators or barriers to guideline implementation, such as cost-effectiveness or changes to organizational culture, is significantly more challenging, particularly when developing international guidelines. This difficulty has previously been used as an explanation for the replication of guidelines when high-quality international guidelines already exist. However, this study debunked that theory for IBD guidelines, as single-country guidelines still scored poorly within this domain.

Since AGREE-S was developed recently as an extension of AGREE-II for surgical guidelines, we examined whether the three new/modified items influenced appraisal results. The items unique to AGREE-S (item 3 in domain 3, item 5 in domain 3 and item 4 in domain 4) had slightly higher variability compared with other items, suggesting some challenges in interpreting expectations for surgical guideline evidence evaluation and recommendation development. However, this did not have a substantial impact on overall domain or guideline scores. Further user experiences with AGREE-S will help refine its specific utility for surgical guideline appraisal.

The use of GRADE has become the gold-standard methodological approach for the transparent development of clinical practice

guidelines, and is used by major healthcare organizations such as the Cochrane Collaboration and the World Health Organization [36-38]. The quality of surgical research is a hotly debated topic, with data on outcomes following surgical intervention being scarce and widely considered poor quality, especially when compared with research on medical interventions [39]. This issue was highlighted in both ECCO guidelines, which stated that GRADE methodology could not be used as the basis for their surgical recommendations due to the poor quality of evidence available. Instead, they used the CEBM tool, one of six grading systems evaluated by the GRADE working group when developing GRADE methodology to address the shortcomings of existing systems used to grade levels of evidence and strength of recommendations [40,41].

The use of alternative methodology to GRADE does give rise to an important question. If the data are considered not to be of sufficient quality to support the use of GRADE methodology, should they serve as the basis for guideline recommendations that can significantly impact the delivery of patient care? The AGREE-S tool does not directly comment on the use of the correct methodology or that the methodology has been clearly described. It does, however, score for whether there has been an explicit link between the synthesis of the evidence and the formulated recommendation, which is therefore where the non-GRADE guidelines would score poorly.

For topics where there is little to no good-quality evidence, expert opinion is usually sought to allow the formulation of recommendations to help guide decision and policy making. It is this lack of good-quality evidence and the need for expert opinion that may have led to the differing recommendations between the guidelines, especially those recommendations surrounding operative techniques.

Recommendations based on expert opinion are usually best coordinated by formal consensus methods [42]. Consensus methods can help to overcome the challenges of gathering opinions from a small group, and the use of formal consensus methodology is recognized as being more reliable than individual expert opinions [42-45]. It is notable that 11 of the 35 papers initially identified were consensus statements rather than true guidelines, possibly indicating a paucity of high-quality evidence in this field. The high proportion of expert consensus highlights the importance of using rigorous methodology, such as formal techniques like Delphi, to optimize the validity of recommendations in the absence of robust scientific evidence.

With the development of the ACCORD (ACcurate CONsensus Reporting Document) tool in early 2024, the quality of consensus based reporting can now be appraised, with the rigour of the consensus methods used to guide recommendations being transparent and clear for readers [42]. Similar to the improvements in guideline quality brought about by development of the AGREE tool, there may be a similar improvement in the quality of surgical consensus statements, thus positively influencing the perception of recommendations developed using this methodology. With the ACCORD tool offering a checklist to improve the reporting and quality of consensus statements, high-quality consensus statements should be developed for clinical research questions where recommendations cannot be made using GRADE methodology due to the poor quality of available evidence.

LIMITATIONS

A limitation of this study is the subjectivity of the AGREE-S tool, which is acknowledged by other studies reviewing guidelines in different medical and surgical conditions [46]. Various factors contribute to the subjectivity of the tool, including the varying experience of reviewers and the direct comparison of papers with one another (as most reviewers appraised multiple papers in the same sitting). Those developing guidelines should ensure that appropriate methodology has been followed as well as the journals publishing the guidelines through the involvement of methodologists in the peer-review process. In this way, less responsibility would fall upon the reader to fully understand guideline methodology and carry out rigorous appraisal.

One guideline was appraised by four reviewers rather than the five reviewers that other guidelines were subject to; this was due to the fifth reviewer being the primary author of the guideline in question, and therefore they did not score their own work to avoid a conflict of interest.

Within the guidelines themselves, we note that guideline development was often poorly or confusingly documented, with some information only available in supplement documents that were not easily found. This may have made it challenging for reviewers to provide a representative score of the true quality of the guideline methodology.

It is important to note that while AGREE-S is a valuable tool for assessing guideline quality, it does not directly account for country-specific factors such as healthcare system structure or medication availability that can influence local implementability. Guideline users should consider these factors when deciding how to apply recommendations in their specific context.

Lastly, this scoping review only included guidelines in the English language.

CONCLUSION

This review demonstrates the breadth of available guidelines for the surgical management of IBD in adults. When scored according to AGREE-S there is great variability in the quality of published guidelines. It is therefore essential to encourage future development committees to use the AGREE-S checklist when formulating new guidelines, to ensure that they are clear, user-friendly and methodologically robust. We also highlight the need for high-quality, collaborative, international guidelines that can be adopted by the healthcare systems of many countries, rather than the replication of de novo guidelines by individual countries [34].

AUTHOR CONTRIBUTIONS

Zarnigar Mussarat Khan: Conceptualization; investigation; writing – original draft; writing – review and editing; methodology; validation; formal analysis; project administration; data curation; visualization. **Camille Ball:** Methodology; conceptualization. **Dalha Saeed:**

Conceptualization; methodology. **Grace Tai:** Conceptualization; methodology. **Shaneil Chandran:** Investigation. **Abhishek Vashista:** Investigation. **Simon Davey:** Investigation; writing – review and editing. **Matthew James Lee:** Investigation; writing – review and editing. **Steven R. Brown:** Investigation; writing – review and editing. **Daniel Hind:** Methodology; writing – original draft; investigation; writing – review and editing. **Adele Elizabeth Sayers:** Conceptualization; investigation; methodology; writing – original draft; validation; writing – review and editing; formal analysis; project administration; supervision; data curation.

FUNDING INFORMATION

No funding was received for this scoping review. Where mentioned, the funders of the individual included guidelines are listed within the original guideline.

CONFLICT OF INTEREST STATEMENT

SB is an author of one of the papers reviewed. To minimize conflict of interest, he did not critically appraise this paper.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ETHICS STATEMENT

None required.

ORCID

Zarnigar Mussarat Khan  <https://orcid.org/0000-0003-3145-2221>

Steven R. Brown  <https://orcid.org/0000-0002-0980-2793>

REFERENCES

1. Wang R, Li Z, Liu S, Zhang D. Global, regional and national burden of inflammatory bowel disease in 204 countries and territories from 1990 to 2019: a systematic analysis based on the global burden of disease study 2019. *BMJ Open*. 2023;13(3):e065186 Available from: <https://bmjopen.bmj.com/content/bmjopen/13/3/e065186.full.pdf>
2. McDowell C, Farooq U, Haseeb M. Inflammatory bowel disease. *StatPearls*. 2023 [cited 2023 Sep 22]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK470312/>
3. Guerra-Farfan E, Garcia-Sanchez Y, Jornet-Gibert M, Nuñez JH, Balaguer-Castro M, Madden K. Clinical practice guidelines: the good, the bad, and the ugly. *Injury*. 2023;1(54):S26–S29.
4. Murad MH. Clinical Practice Guideline. *Mayo Clinic Proc*. 2017;92:423–33. Available from: <http://www.mayoclinicproceedings.org/article/S0025619617300253/fulltext>
5. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *Br Med J*. 1999;318:527–30. Available from: <https://pmc/articles/PMC1114973/>
6. Zhang Y, Akl EA, Schünemann HJ. Using systematic reviews in guideline development: the GRADE approach. *Res Synth Methods*. 2018. Available from: <https://doi.org/10.1002/jrsm.1313>. Epub ahead of print.
7. Institute of Medicine (US) Committee on Standards for Developing Trustworthy Clinical Practice Guidelines, Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E, editors. *Clinical practice*

- guidelines we can trust. Washington (DC): National Academies Press; 2011. Available from: <https://doi.org/10.17226/13058>
8. Martínez García L, Sanabria AJ, García Álvarez E, Trujillo-Martín MM, Etxeandia-Ikobaltzeta I, Kotzeva A, et al. The validity of recommendations from clinical guidelines: a survival analysis. *CMAJ*. 2014;186(16):1211–9.
 9. Kung J, Miller RR, Mackowiak PA. Failure of clinical practice guidelines to meet Institute of Medicine Standards: two more decades of little, if any, progress. *Arch Intern Med*. 2012;172(21):1628–33. Available from: <https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/1384245>
 10. Brouwers MC, Kerkvliet K, Spithof K. AGREE Next Steps Consortium. The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines. *BMJ*. 2016;352:i1152. Available from: <https://10.1136/bmj.i1152>
 11. Antoniou SA, Florez ID, Markar S, Logullo P, López-Cano M, Silecchia G, et al. AGREE-S: AGREE II extension for surgical interventions: appraisal instrument. *Surg Endosc*. 2022;36(8):5547–58. <https://doi.org/10.1007/s00464-022-09354-z>
 12. Zambrano-Sánchez R, Alvarez-Mena P, Hidalgo D, Liquitay CME, Franco JVA, Vernooij RWM, et al. Quality assessment of clinical practice guidelines (CPG) for the diagnosis and treatment of inflammatory bowel disease using the AGREE II instrument: a systematic review. *BMC Gastroenterol*. 2022;22(1):1–14. Available from: <https://bmcgastroenterol.biomedcentral.com/articles/10.1186/s12876-022-02539-9>
 13. Haddaway NR, Collins AM, Coughlin D, Kirk S. The role of Google Scholar in evidence reviews and its applicability to grey literature searching. *PLoS One*. 2015;10(9):e0138237 Available from: <http://www.bangor.ac.uk/>
 14. The Gap Consortium. AGREE-S appraisal instrument—AGREE-S. [cited 2023 Sep 26]. Available from: <https://agree-s.org/agree-s-appraisal-instrument/>
 15. Adamina M, Bonovas S, Raine T, Spinelli A, Warusavitarne J, Armuzzi A, et al. ECCO guidelines on therapeutics in Crohn's disease: surgical treatment. *J Crohn's Colitis*. 2020;14(2):155–68. Available from: <https://moh-it.pure.elsevier.com/en/publications/ecco-guidelines-on-therapeutics-in-crohns-disease-surgical-treatm>
 16. Bernstein CN, Eliakim A, Fedail S, Fried M, Gearry R, Goh KL, et al. World Gastroenterology Organisation global guidelines inflammatory bowel disease. *J Clin Gastroenterol*. 2016;50(10):813–8.
 17. Brown SR, Fearnhead NS, Faiz OD, Abercrombie JF, Acheson AG, Arnott RG, et al. The Association of Coloproctology of Great Britain and Ireland consensus guidelines in surgery for inflammatory bowel disease. *Colorectal Dis*. 2018;20:3–117.
 18. Eder P, Łodyga M, Gawron-Kiszka M, Dobrowolska A, Gonciarz M, Hartleb M, et al. Guidelines for the management of ulcerative colitis. Recommendations of the Polish Society of Gastroenterology and the Polish National Consultant in Gastroenterology. *Prz Gastroenterol*. 2023;18(1):1–42. <https://doi.org/10.5114/pg.2023.125882>
 19. Eliadou E, Day AS, Thompson-Fawcett MW, Gearry RB, Rowbotham DS, Walmsley R, et al. New Zealand Society of Gastroenterology guidelines for the management of refractory ulcerative colitis. *NZ Med J*. 2015;128(1423):63–76.
 20. Holubar SD, Lightner AL, Poylin V, Vogel JD, Gaertner W, Davis B, et al. The American Society of Colon and Rectal Surgeons clinical practice guidelines for the surgical management of ulcerative colitis. *Dis Colon Rectum*. 2021;64(7):783–804.
 21. Kucharzik T. Updated S3-guideline ulcerative colitis. German Society for Digestive and Metabolic Diseases (DGVS). *Z Gastroenterol*. 2019;57(2):162–241. <https://doi.org/10.1055/a-0824-0861>
 22. Lightner AL, Vogel JD, Carmichael JC, Keller DS, Shah SA, Mahadevan U, et al. The American Society of Colon and Rectal Surgeons clinical practice guidelines for the surgical management of Crohn's disease. *Dis Colon Rectum*. 2020;63(8):1028–52.
 23. Łodyga M, Eder P, Gawron-Kiszka M, Dobrowolska A, Gonciarz M, Hartleb M, et al. Guidelines for the management of patients with Crohn's disease. Recommendations of the Polish Society of Gastroenterology and the Polish National Consultant in Gastroenterology. *Prz Gastroenterol*. 2021;16(4):257–96.
 24. Matsuoka K, Kobayashi T, Ueno F, Matsui T, Hirai F, Inoue N, et al. Evidence-based clinical practice guidelines for inflammatory bowel disease. *J Gastroenterol*. 2018;53(3):305–53. <https://doi.org/10.1007/s00535-018-1439-1>
 25. Mowat C, Cole A, Windsor A, Ahmad T, Arnott I, Driscoll R, et al. Guidelines for the management of inflammatory bowel disease in adults. *Gut*. 2011;60(5):571–607.
 26. Park JJ, Yang SK, Ye BD, Kim JW, Il PD, Yoon H, et al. Second Korean guidelines for the management of Crohn's disease. *Intestinal Res*. 2017;15:38–67. <https://doi.org/10.5217/ir.2017.15.1.38>
 27. Spinelli A, Bonovas S, Burisch J, Kucharzik T, Adamina M, Annese V, et al. ECCO guidelines on therapeutics in ulcerative colitis: surgical treatment. *J Crohn's Colitis*. 2022;16(2):179–89. Available from: <https://academic.oup.com/ecco-jcc/article/16/2/179/6390023>
 28. Strong S, Steele SR, Boutsous M, Bordineau L, Chun J, Stewart DB, et al. Clinical practice guideline for the surgical management of Crohn's disease. *Dis Colon Rectum*. 2015;58(11):1021–36. Available from: <https://journals.lww.com/00003453-201511000-00001>
 29. Ueno F, Matsui T, Matsumoto T, Matsuoka K, Watanabe M, Hibi T. Evidence-based clinical practice guidelines for Crohn's disease, integrated with formal consensus of experts in Japan. *J Gastroenterol*. 2013;48:31–72. <https://doi.org/10.1007/s00535-012-0673-1>
 30. Siemieniuk R, Gordon G. What is GRADE? *BMJ Best Pract*. 2018;47:1–5. Available from: <https://bestpractice.bmj.com/info/toolkit/learn-ebm/what-is-grade/>
 31. Nuffield Department of Primary Care Health Sciences (Medical Sciences Division). Critical appraisal tools—Centre for Evidence-Based Medicine (CEBM). Oxford, UK: University of Oxford; 2020 [cited 2024 Mar 29]. Available from: <https://www.cebm.ox.ac.uk/resources/ebm-tools/critical-appraisal-tools>
 32. Tovey D. How to clarify a clinical question. *BMJ Best Pract*. 2020;1–3. [cited 2024 Mar 29]. Available from: <https://bestpractice.bmj.com/info/toolkit/learn-ebm/how-to-clarify-a-clinical-question/>
 33. Armstrong JJ, Goldfarb AM, Instrum RS, MacDermid JC. Improvement evident but still necessary in clinical practice guideline quality: a systematic review. *J Clin Epidemiol*. 2017;81:13–21. Available from: <https://pubmed.ncbi.nlm.nih.gov/27565978/>
 34. McAllister M, Florez ID, Stoker S, McCaul M. Advancing guideline quality through country-wide and regional quality assessment of CPGs using AGREE: a scoping review. *BMC Med Res Methodol*. 2023;23(1):1–12. Available from: <https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-023-02101-5>
 35. Guidelines International Network. Home Page—GIN. 2024 [cited 2024 Jul 25]. Available from: <https://g-i-n.net/>
 36. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–6. Available from: <https://www.bmj.com/content/336/7650/924>
 37. Schünemann H, Oxman JBGGA. GRADE handbook. 2013 [cited 2024 Jun 14]. Available from: <https://gdt.gradepro.org/app/handbook/handbook.html>
 38. World Health Organization WHO guideline on school health services. Web Annex F. Systematic reviews of the effectiveness and acceptability of comprehensive school health services: GRADE evidence profiles and evidence-to-decision table. 2021 [cited 2024 Jun 14]. Available from: <http://apps.who.int/>
 39. Domenghino A, Walbert C, Birrer DL, Pahan MA, Clavien PA, Outcome4Medicine consensus group. Consensus recommendations on how to assess the quality of surgical interventions. *Nat Med*. 2023;29:811–22. Available from: <https://10.1038/s41591-023-02237-3>

40. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, et al. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches. *BMC Health Serv Res.* 2004;4(1):38 Available from: <https://pubmed.ncbi.nlm.nih.gov/15615589/>
41. ECDC Technical. Evidence-based methodologies for public health. 2011 [cited 2024 Jun 15]. Available from: www.ecdc.europa.eu
42. Gattrell WT, Logullo P, van Zuuren EJ, Price A, Hughes EL, Blazey P, et al. ACCORD (ACcurate COnsensus reporting document): a reporting guideline for consensus methods in biomedicine developed via a modified Delphi. *PLoS Med.* 2024;21(1):e1004326 Available from: <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1004326>
43. Kurvers RHJM, Herzog SM, Hertwig R, Krause J, Carney PA, Bogart A, et al. Boosting medical diagnostics by pooling independent judgments. *Proc Natl Acad Sci USA.* 2016;113(31):8777–82. Available from: <https://pubmed.ncbi.nlm.nih.gov/27432950/>
44. Surowiecki J. *The wisdom of crowds.* New York City, USA: Anchor; 2004.
45. Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW. Evidence for a collective intelligence factor in the performance of human groups. *Science.* 2010;330(6004):686–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/20929725/>
46. Hoffmann-Eßer W, Siering U, Neugebauer EAM, Brockhaus AC, McGauran N, Eikermann M. Guideline appraisal with AGREE II: on-line survey of the potential influence of AGREE II items on overall assessment of guideline quality and recommendation for use. *BMC Health Serv Res.* 2018;18(1):143.

How to cite this article: Khan ZM, Ball C, Saeed D, Tai G, Chandran S, Vashista A, et al. Appraisal of current surgical guidelines for inflammatory bowel disease using the AGREE-S instrument: A scoping review. *Colorectal Dis.* 2025;27:e17258. <https://doi.org/10.1111/codi.17258>