



Appraisal of the current guidelines for the management of diverticular disease using the Appraisal of Guidelines Research and Evaluation II (AGREE II) instrument

P Gavriilidis¹, A Askari², E Gavriilidis³, N de'Angelis^{4,5}, S Di Saverio⁶, J Wheeler², RJ Davies²

¹Imperial College Healthcare NHS Trust, Hammersmith Hospital, UK

²Cambridge University Hospitals NHS Foundation Trust, UK

³University Hospital of Alexandroupolis, Democritus University of Thrace, Alexandroupolis, Greece

⁴University Hospital Henri Mondor (AP-HP), Créteil, France

⁵University of Paris Est, Créteil, France

⁶University of Insubria, Insubria, Italy

ABSTRACT

Introduction Diverticular disease is one of the most frequent reasons for attending emergency departments and surgical causes of hospital admission. In the past decade, many surgical and gastroenterological societies have published guidelines for the management of diverticular disease. The aim of the present study was to appraise the methodological quality of these guidelines using the Appraisal of Guidelines Research and Evaluation II (AGREE II) tool.

Methods PubMed, Embase, Cochrane Library and Google Scholar databases were searched systematically. The methodological quality of the guidelines was appraised independently by five appraisers using the AGREE II instrument.

Findings A systematic search of the literature identified 12 guidelines. The median overall score of all guidelines was 68%. Across all guidelines, the highest score of 85% was demonstrated in the domain 'Scope and purpose'. The domains 'Clarity and presentation' and 'Editorial independence' both scored a median of 72%. The lowest scores were demonstrated in the domains 'Stakeholder involvement' and 'Applicability' at 46% and 40%, respectively. Overall, the National Institute for Health and Care Excellence (NICE) guidelines performed consistently well, scoring 100% in five of six domains; NICE was one of the few guidelines that specifically reported stakeholder involvement, scoring 97%. Generally, the domain of 'Stakeholder involvement' ranked poorly with seven of twelve guidelines scoring below 50%, with the worst score in this domain demonstrated by Danish guidelines at 25%.

Conclusion Six of twelve guidelines (NICE, American Society of Colon & Rectal Surgeons (ASCRS), European Society of Coloproctology (ESCP), American Gastroenterological Association, German Society of Gastroenterology/German Society for General and Visceral Surgery (German), Netherlands Society of Surgery) scored above 70%. Only three, NICE, ASCRS and ESCP, scored above 75% and were voted unanimously by the appraisers for use as they are. Therefore, use of AGREE II may help improve the methodological quality of guidelines and their future updates.

KEYWORDS

Clinical practice guidelines – Appraisal of Guidelines Research and Evaluation II – Diverticular disease – Diverticulitis

Accepted 30 December 2020

CORRESPONDENCE TO

Paschalis Gavriilidis, E: pgavriilidis@yahoo.com

Introduction

Diverticular disease is a common gastrointestinal disorder in industrialised countries. It has been reported that the prevalence of diverticulosis in patients above 60 years of age is approximately 55%. However, in recent decades, the incidence rate of diverticulitis has increased by 132% in those aged 40–49 years. The prevalence of emergency hospital admission for diverticulitis varies widely between different ethnic groups. In the USA, for example, for Caucasian patients the incidence appears to be 62 in

100,000. This rate is somewhat lower in Hispanic and African American patients (50 in 100,000) and lower still for patients from East Asia (10 in 100,000). Furthermore, the annual cost of diverticulitis to healthcare providers has been estimated at over \$2 billion.^{1–3}

In the past decade, 12 clinical practice guidelines for the management of diverticular disease have been published.^{4–15} In 2018, Galetin *et al* reported that major discordances were present between the guidelines across a wide variety of issues, including classification of disease,

use of computed tomography versus ultrasonography, need for antibiotics in outpatient treatment, and mode of surgery. The available evidence presented at the time, however, was rated as moderate or low quality.¹⁶

More than 40 tools have been developed for appraisal of the methodological quality of clinical practice guidelines; the most recent tool is the Appraisal of Guidelines Research and Evaluation II (AGREE II) instrument which is validated internationally and supported by the World Health Organization (WHO) Advisory Committee on Health Research, and by many guideline development teams.^{17,18} Further information is available at www.agreetrust.org.¹⁹ The aim of the present study was to appraise the methodological quality of clinical practice guidelines for the management of diverticular disease using the AGREE II instrument.

Methods

Search strategy and guidelines selection

A systematic literature search was conducted independently by two of the authors (PG and AA) using PubMed, Embase, Cochrane Library and Google Scholar, and the terms clinical practice guidelines, diverticular disease, diverticulosis, diverticulitis, colonic and obstruction. The search was limited to the past 10 years and guidelines published in the English language. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) was used to demonstrate the steps of the search strategy and the rationale for any exclusions.²⁰ After independent evaluation of guidelines by two authors (PG and AA) the following data were extracted: country of origin, year of publication, developers, founding resources and evaluation measures.

Appraisal of guidelines

The AGREE II tool was used to assess the quality of the guidelines. The checklist comprises 23 items divided into six domains: Scope and purpose, Stakeholder involvement, Rigour of development, Clarity of presentation, Applicability, and Editorial independence. For further details regarding the criteria used to describe and evaluate the six domains and the 23 constituent items, see Supplementary Table 1. After undergoing online training (www.agreetrust.org) to ensure appraisal standardisation, five appraisers (PG, AA, EG, NDA, SDS), as recommended by the AGREE II consortium, evaluated the guidelines independently using the AGREE II tool (September 2013 version). As per the AGREE II manual, discrepancies of more than two standard deviations (SD) were resolved through discussion.

Findings

Search strategy and guidelines selection

Search results using the keywords: clinical practice guidelines, diverticular disease, diverticulosis, diverticulitis,

colonic and obstruction returned a total of 423 articles, ten of which were duplicates. Of the remaining 413 abstracts screened, 398 were excluded because they were not guidelines, leaving 15 guidelines for examination. Of these, three were excluded because they were previous versions of updated guidelines, resulting in a total of 12 guidelines included in this study (Figure 1). The studies originated from a variety of countries and regions, and some were the result of international collaboration. The guidelines used a variety of scoring systems in assessing the quality of evidence including the Oxford, Delphi and Grade of Recommendations Assessment, Development and Evaluation (GRADE) systems (Table 1).

AGREE II appraisal

The median overall score across all guidelines was 68%. The highest score was achieved by the National Institute for Health and Care Excellence (NICE) guidelines (100%) and the lowest by the Polish guidelines (26%). The highest score across all the guidelines was observed in domain I (Scope and purpose) (85%), followed by domain IV (Clarity of presentation) and domain VI (Editorial independence), both scoring a median of 72%. The worst scoring domains were V (Applicability) and II (Stakeholder involvement), scoring a median of 40% and 46% respectively (Figure 2).

In domain I (Scope and purpose), the scores were generally high (median 85%). In particular, second best after the NICE with a score of 100% was the American Society of Colon & Rectal Surgeons (ASCRS) with a score of 89%; the worst score was observed in the Polish guidelines at 47%.

In domain II (Stakeholder involvement) the highest score (97%) was observed in NICE and the lowest (25%) in the Danish guidelines.

In domain III (Rigour of development), NICE guidelines scored 100%, followed by ASCRS with a score of 79%; the lowest score was observed in American Association of Family Physicians guidelines.

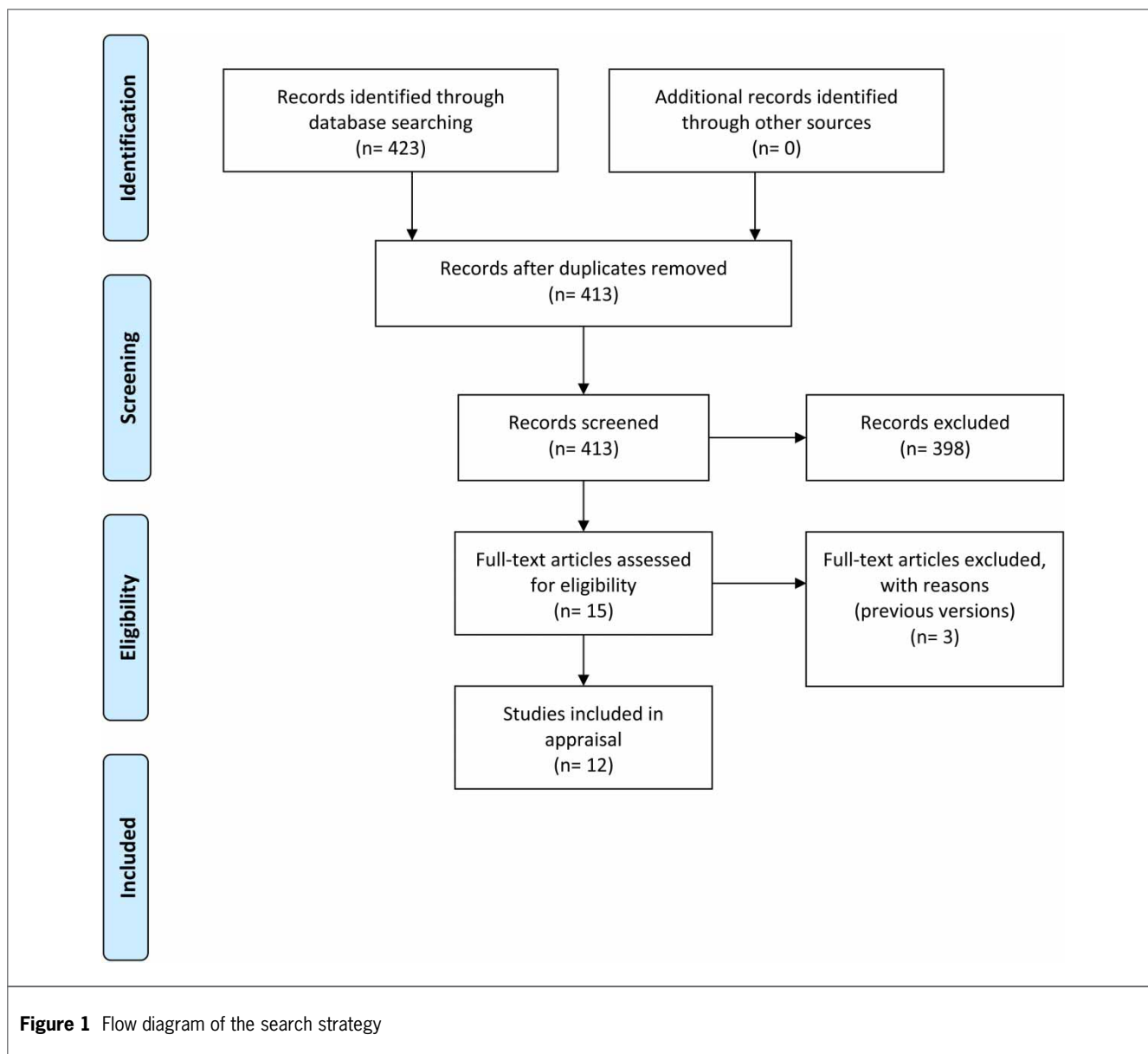
In domain IV (Clarity of presentation), NICE scored 100%, followed by the German Society of Gastroenterology/German Society for General and Visceral Surgery (German) guidelines with a score of 92%; the lowest score was observed in the Polish guidelines (47%).

In domain V (Applicability), NICE scored 100%, followed by the German guidelines with a score of 54%; the lowest score was observed in the Polish guidelines (23%).

In domain VI (Editorial independence), NICE scored 100% with second place shared by three guidelines, namely American Gastroenterological Association (AGA), ASCRS and World Society of Emergency Surgery, each with a score of 92%. The worst score was observed in Danish guidelines at 0% (Figure 3).

Overall scores of each guideline

Overall, the NICE guidelines performed consistently well, scoring 100% in five of the six domains. Second place was shared by two guidelines, ASCRS and European Society of Coloproctology (ESCP), both of which scored 79%. Third place was shared by three guidelines, AGA, German and



Netherlands Society of Surgery (NSS), which all scored 71%. Six guidelines scored more than 70% and were considered of high quality (AGA, ASCRS, ESCP, German, NICE and NSS). Three guidelines scored more than 75% (ASCRS, ESCP, NICE) and were voted unanimously by the appraisers for use as they are (Figure 4).

Discussion

This study appraises for the first time the methodological quality of guidelines for diverticular disease using the AGREE II instrument. The overall quality of the guidelines was good with a median score of 68%, and in particular, six of twelve guidelines scored above 70% and

can be considered of high quality (Figure 4). However, similar appraisals of guidelines from hepatobiliary surgery demonstrated worse performance compared with diverticular disease guidelines.^{21,22}

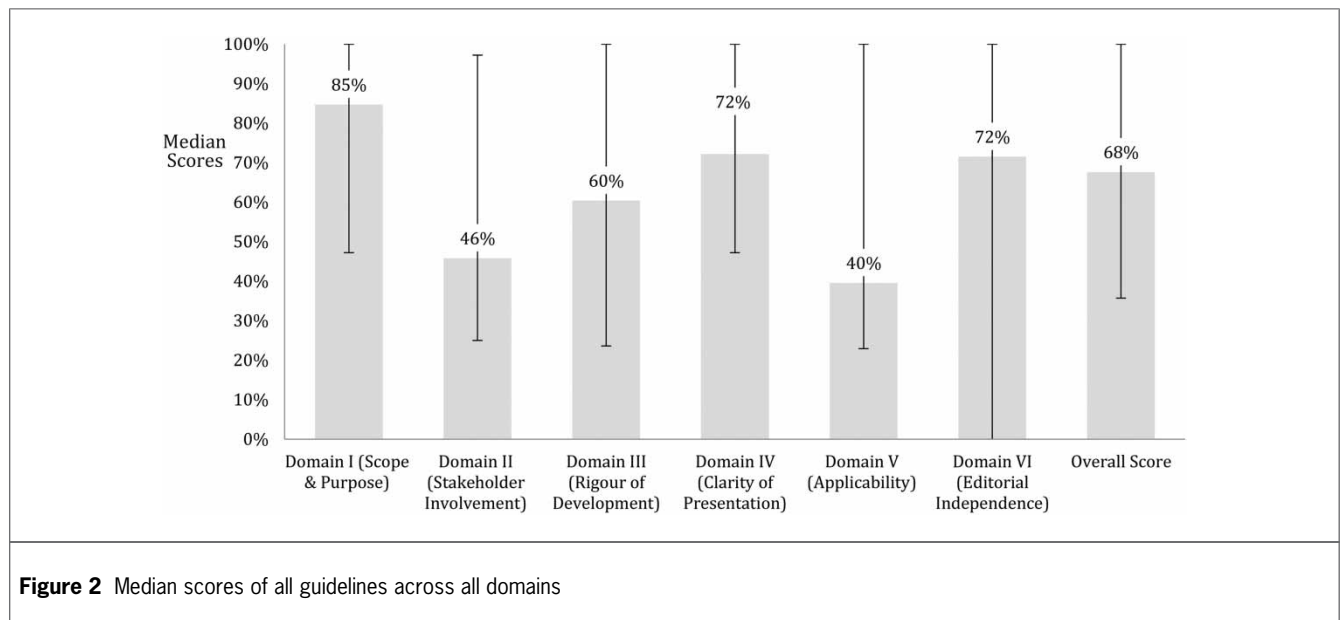
It has been reported that the methodological quality of guidelines across all specialties is improving over time. In particular, guidelines usually score well in the 'Scope and purpose' domain.¹⁸

In the present study, in the domain 'Scope and purpose', all guidelines bar one scored above 54%. The scope and purpose of guidelines had been stated and described clearly by the stakeholders. However, in the domain 'Stakeholder involvement', where the involvement of experts of all specialties secures important attributes and insights and promotes constructive debates between the

Table 1 Study characteristics

Developer, region, year	Content	Scoring system	
NICE, UK, 2020	E, D&T, FU	NICE system	Focus on diverticular disease
ESCP, Europe, 2020	E, D&T, FU	Oxford	Focus on diverticular disease
WSES, World, 2020	E, D&T, FU	GRADE	Focus on acute left-sided diverticulitis
ASCRS, USA, 2020	E, D&T, FU	GRADE	Focus on diverticular disease
Polish, Poland, 2015	E, D&T, FU	Do Not Grade the Strength	Focus on Diverticular Disease
SICCR, Italy, 2015	E, D&T, FU	GRADE	Focus on diverticular disease
GSG/GSGVS, Germany, 2014	E, D&T, FU	DCCG system Delphi process	Focus on diverticular disease
AGA, USA, 2015	E, D&T, FU	GRADE	Focus on acute left-sided diverticulitis
NSS, the Netherlands, 2013	E, D&T, FU	Oxford	Focus on acute left-sided diverticulitis
AAFP, USA, 2013	E, D&T, FU	SORT	Focus on acute left-sided diverticulitis
EAES, Europe, 2012	E, D&T, FU	Oxford Delphi	Focus on surgical aspects
DCCG, Denmark, 2012	E, D&T, FU	DCCG system	Focus on diverticular disease

NICE = National Institute for Health and Care Excellence; ESCP = European Society of Coloproctology; WSES = World Society of Emergency Surgery; GRADE = Grade of Recommendations Assessment, Development and Evaluation; ASCRS = American Society of Colon & Rectal Surgeons; SICCR = Italian Society of Colon and Rectal Surgeons; GSG/GSGVS = German Society of Gastroenterology/German Society for General and Visceral Surgery; AGA = American Gastroenterological Association; NSS, Netherlands Society of Surgery; AAFP = American Association of Family Physicians; EAES = European Association of Endoscopic Surgery; DCCG = Danish Surgical Society; E = epidemiology; D&T = diagnosis & treatment; FU = follow-up



included experts, there was misrepresentation of experts of all related specialties. Furthermore, there was misrepresentation of patients to whom the guidelines apply. It has been reported that patient participation in the formation of guidelines reduces the risk of adopting unhelpful recommendations that may have been included by certain stakeholders with potential self-serving

interests.²⁵ Unfortunately, generally poor scores were observed in the domain ‘Stakeholder involvement’. In particular, the highest score was achieved by NICE at 97% and the lowest by the Danish guidelines at 25%.

Another area of guidelines that could be considered the Achilles’ heel is the domain ‘Applicability’. It has been reported by previous AGREE assessments that guidelines

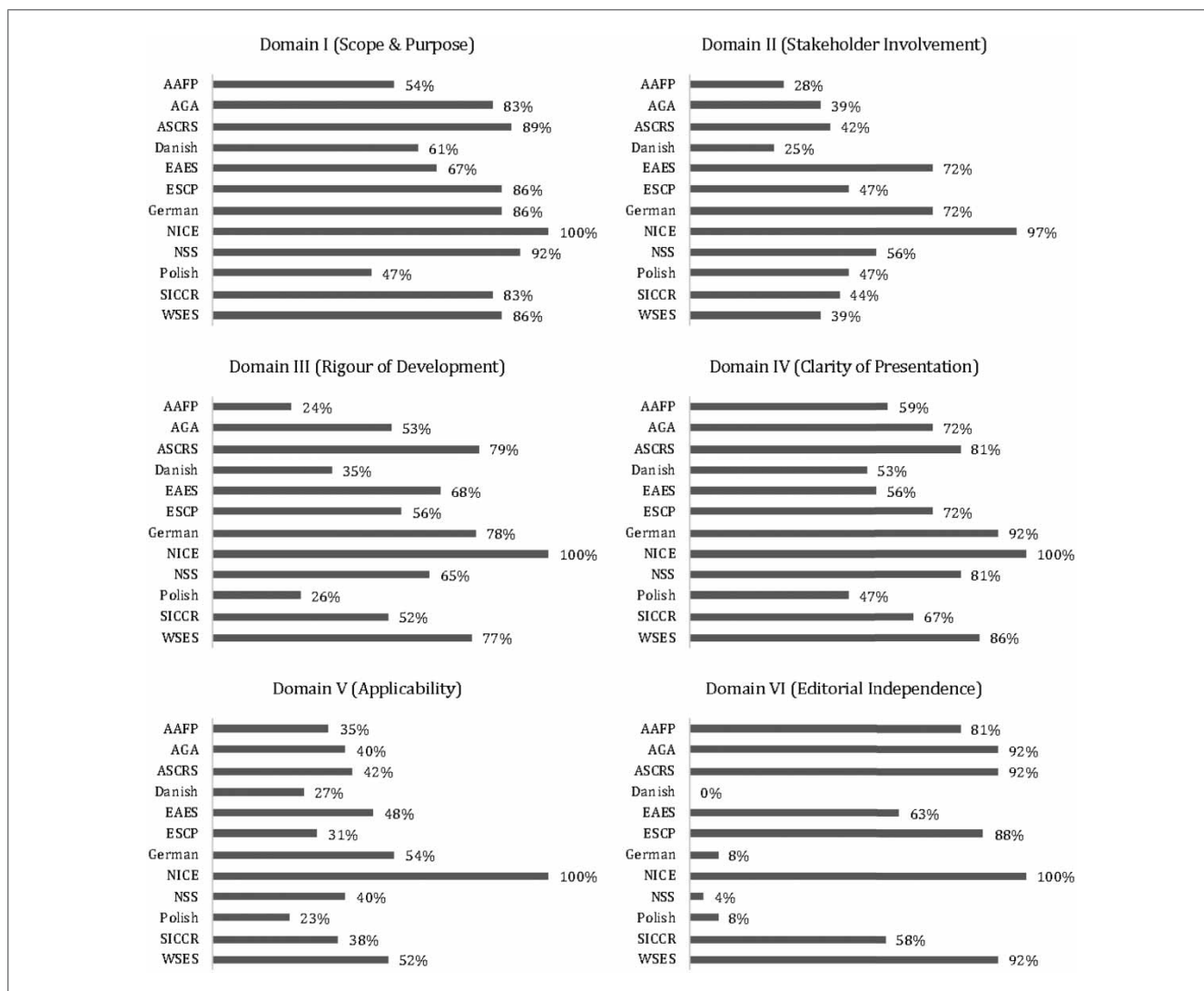


Figure 3 Individual domain scores for each guideline. AAFF = American Association of Family Physicians; AGA = American Gastroenterological Society; ASCRS = American Society of Colon & Rectal Surgeons; EAES = European Association of Endoscopic Surgery; ESCP = European Society of Coloproctology; NICE = National Institute for Health and Care Excellence; NSS = Netherlands Society of Surgery; SICCR = Italian Society of Colorectal Surgery; WSES = World Society of Emergency.

lack helpful instructions for their practical application.²⁴ In the present study, this weakness was demonstrated in several of the guidelines with a score of 40%, the worst median score across the six domains (Figure 2).

In the domain ‘Rigour and development’ where search methods, evidence selection criteria, strengths and limitations of evidence, formulation of recommendations and considerations of benefits and harms were evaluated, the guidelines performed fairly well, scoring 60%. In particular, nine of twelve guidelines scored above 52% (Figure 3). This could be due to the universal use of validated appraisal systems such as GRADE and Oxford.

Furthermore, it has been reported that domain III ‘Rigour and development’ and domain V ‘Applicability’

had a special impact on the decision by the clinician appraiser of whether to use the guideline or not.²⁵ Domain III provides guidance for weighting the specific guideline quality on whether systematic methods were used to search for evidence. Clear description of the selecting criteria, the strengths and limitations of the body of evidence, the methods of formulating the recommendations and whether the guideline has been externally reviewed by a group of experts were the principal prerequisites for a high score. A high score in domain III usually demonstrates guidelines based on high-quality evidence and minimum bias.²⁶ On the other hand, a low score demonstrates either a lack of methodological expertise by the developers of the

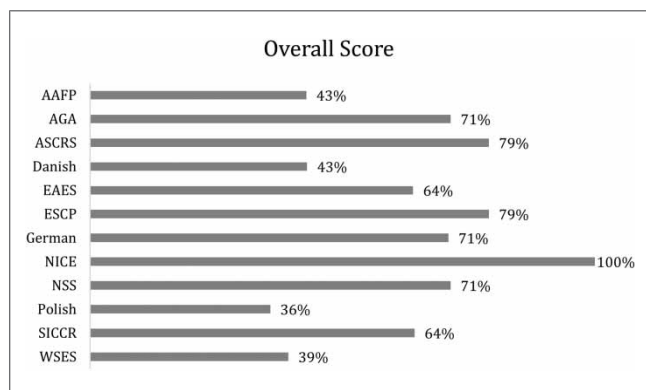


Figure 4 Overall scores for each guideline. AAFP = American Association of Family Physicians; AGA = American Gastroenterological Society; ASCRS = American Society of Colon & Rectal Surgeons; EAES = European Association of Endoscopic Surgery; ESCP = European Society of Coloproctology; NICE = National Institute for Health and Care Excellence; NSS = Netherlands Society of Surgery; SICCR = Italian Society of Colorectal Surgery; WSES = World Society of Emergency

guideline or an insufficient systematic search of the literature due to lack of resources.¹⁸ Furthermore, the waste of resources could be enhanced if a guideline is implemented insufficiently by clinicians in everyday practice because of inadequate implementation instructions. This may lead to omission of beneficial therapeutic strategies, and less ability to prevent harm.²⁷

This evidence underlines the importance of domain V (Applicability). An important weakness that should be underlined regarding the domain of applicability is that none of the questions in this domain refer to the specific applicability of the guideline. It has been reported that a guideline may achieve a high score even though it is not applicable to a specific patient population.²⁸ A future edition of AGREE II should address this weakness. Furthermore, AGREE II has been criticised due to lack of specific questions on how to evaluate the domain of the overall assessment and in view of this some researchers suggest that these assessments are subjective.^{29,30}

Reflecting on the above evidence we would like to underline that our study should be read in the context of its limitations, some of which can be attributed to the very nature of the AGREE II system. For example, the AGREE II checklist has been criticised for its assumption that all domains carry the same weight in terms of scores. Another potential source of positive bias could have been the influence that may have been shown to guidelines produced by reputable international societies and institutions. Conversely, guidelines produced by societies with a 'lack of international prestige' may not be scored as highly. This caveat is a consequence of non-blinded assessment.

Conclusion

This first appraisal of guidelines for diverticular disease using the AGREE II instrument demonstrated that six of the twelve guidelines are of high quality. Furthermore, the appraisers voted unanimously for use of NICE, ASCRS and ESCP guidelines as they are. Therefore, future guideline updates considering the limitations highlighted by the AGREE II appraisal might lead to improved guideline quality.

Conflicts of interest

None.

Funding

None.

Author contributions

PG: study concept and design; acquisition of data; analysis and interpretation of data; drafting the manuscript; appraisal the guidelines; statistical analysis; critical revision of the manuscript for important intellectual content. AA: acquisition of data; analysis and interpretation of data; drafting the manuscript; appraisal the guidelines; statistical analysis; critical revision of the manuscript for important intellectual content. EG: analysis and interpretation of data; appraisal the guidelines; critical revision of the manuscript for important intellectual content. NDA: analysis and interpretation of data; appraisal the guidelines; critical revision of the manuscript for important intellectual content. SDS: analysis and interpretation of data; appraisal of the guidelines, drafting the manuscript; critical revision of the manuscript for important intellectual content. JW: analysis and interpretation of data; drafting the manuscript; statistical analysis; critical revision of the manuscript for important intellectual content. RJD: study concept and design; acquisition of data; analysis and interpretation of data; drafting the manuscript; statistical analysis; critical revision of the manuscript for important intellectual content; study supervision.

References

1. Peery AF, Crockett SD, Barrit AS *et al*. Burden of gastrointestinal, liver and pancreatic diseases in the United States. *Gastroenterology* 2015; **149**: 1731–1741.
2. Strate LL, Morris AM. Epidemiology, pathophysiology, and treatment of diverticulitis. *Gastroenterology* 2019; **156**: 1282–1298.
3. Painter NS, Burkitt DP. Diverticular disease of the colon: a deficiency disease of Western civilization. *Br Med J* 1971; **2**(5759): 450–454.
4. National Institute for Health and Care Excellence. *Diverticular Disease: Diagnosis and Management*. NICE guideline NG147. 2019.
5. Schultz JK, Azhar N, Binda GA, *et al*. European society of coloproctology: guidelines for the management of diverticular disease of the colon. *Colorectal Dis* 2020. Online ahead of print.

6. Sartelli M, Weber DG, Kluger Y, et al. Update of the WSES guidelines for the management of acute colonic diverticulitis in the emergency setting. *World J Emerg Surg* 2020; **15**: 32.
7. Hall J, Hardiman K, Lee S, et al. The American society of colon and rectal surgeons clinical practice guidelines for the treatment of left-sided colonic diverticulitis. *Dis Colon Rectum* 2020; **63**: 728–747.
8. Pietrzak A, Bartnik W, Szczepkowski M, et al. Polish interdisciplinary consensus of diagnostic and treatment of colonic diverticulosis. *Pol Przegl Chir* 2015; **87**: 203–220.
9. Binda GA, Cuomo R, Laghi A, et al. Practice parameters for the treatment of colonic diverticular disease: Italian society of colon and rectal surgery (SICCR) guidelines. *Tech Coloproctol* 2015; **19**: 615–626.
10. Kruis W, Germer CT, Leifeld L; German Society for gastroenterology, Digestive and Metabolic Diseases and the German Society for General and Visceral Surgery. Diverticular disease: guidelines of the German society for gastroenterology, digestive and metabolic diseases and the German society for general and visceral surgery. *Digestion* 2014; **90**: 190–207.
11. Stollman N, Smalley W, Hirano I. And AGA institute clinical guidelines committee. American gastroenterological association institute guideline on the management of acute diverticulitis. *Gastroenterology* 2015; **149**: 1944–1949.
12. Andeweg CS, Mulder IM, Felt-Bersma RJF, Verbon A, et al. Guidelines of diagnostics and treatment of acute left-sided colonic diverticulitis. *Dig Surg* 2013; **30**: 278–292.
13. Wilkins T, Embry K, George R. Diagnosis and management of acute diverticulitis. *Am Fam Physician* 2013; **87**: 612–620.
14. Agresta F, Ansaloni L, Baiocchi GL, et al. Laparoscopic approach to acute abdomen from the consensus development conference of the societa Italiana di Chirurgia nell'Ospedality Privata (SICOP), and the European Association for Endoscopic surgery (EAES). *Surg Endosc* 2012; **26**: 2134–2164.
15. Andersen JC, Bundgaard L, Elbrønd H, et al. Danish Surgical Society. Danish national guidelines for treatment of diverticular disease. *Dan Med J* 2012; **59**: C4453.
16. Galetin T, Galetin A, Vestweber KH, Rink AD. Systematic review and comparison of national and international guidelines on diverticular disease. *Int J Colorectal Dis* 2018; **33**: 261–272.
17. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting, and evaluation in health care. *Prev Med* 2010; **51**: 421–424.
18. Alonso-Coello P, Irfan A, Sola I, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. *BMJ Qual Saf* 2010; **19**: e58–e58.
19. Appraisal of Guidelines for Research and Evaluation II, 2013. <http://www.agreetrust.org>.
20. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009; **6**: e1000097.
21. Gavriilidis P, Roberts KJ, Askari A, et al. Evaluation of the current guidelines for resection of hepatocellular carcinoma using the appraisal of guidelines for research and evaluation II instrument. *J Hepatol* 2017; **67**: 991–998.
22. Gavriilidis P, Askari A, Roberts KJ, Sutcliffe RP. Appraisal of the current guidelines for management of cholangiocarcinoma-using the appraisal of guidelines research and evaluation II (AGREE II) instrument. *Hepatobiliary Surg Nutr* 2020; **9**: 126–135.
23. Hämeen-Anttila K, Komulainen J, Enlund H, et al. Incorporating patient perspectives in health technology assessments and clinical practice guidelines. *Res Social Adm Pharm* 2016; **12**: 903–913.
24. O'Donoghue KJM, Reed RD, Knight SR, et al. Systematic review of clinical practice guidelines in kidney transplantation. *BJS Open* 2017; **1**: 97–105.
25. Hoffmann-Eßer W, Siering U, Neugebauer EAM, et al. Guideline appraisal with AGREE II: systematic review of the current evidence on how users handle the 2 overall assessments. *PLoS ONE* 2017; **12**: e0174831.
26. Brosseau L, Rahman P, Poitras S, et al. A systematic critical appraisal of non-pharmacological management of rheumatoid arthritis with appraisal of guidelines for research and evaluation II. *PLoS ONE* 2014; **9**: e95369.
27. Gagliardi AR, Brouwers MC. Do guidelines offer implementation advice to target users? A systematic review of guideline applicability. *BMJ Open* 2015; **5**: e007047.
28. Damiani G, Silvestrini G, Trozzi L, et al. Quality of dementia clinical guidelines and relevance to the care of older people with comorbidity: evidence from the literature. *Clin Interv Aging* 2014; **9**: 1399–1407. Epub 2014/08/30.
29. Lee GY, Yamada J, Kyololo OB, et al. Pediatric clinical practice guidelines for acute procedural pain: a systematic review. *Pediatrics* 2014; **133**: 500–515.
30. Sabharwal S, Patel NK, Gauher S, et al. High methodologic quality but poor applicability: assessment of the AAOS guidelines using the AGREE II instrument. *Clin Orthop* 2014; **472**: 1982–1988.