Review Article

# Appraisal of the current guidelines for management of cholangiocarcinoma—using the Appraisal of Guidelines Research and Evaluation II (AGREE II) Instrument

**Paschalis Gavriilidis[1], Alan Askari[2], Keith J. Roberts[1], Robert P. Sutcliffe[1]**

[1]Department of Hepato-Pancreato-Biliary and Liver Transplant Surgery, Queen Elizabeth University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK; [2]Department of Upper Gastro-Intestinal Surgery, West Hertfordshire Hospitals NHS Trust, Watford, UK

*Contributions:* (I) Conception and design: P Gavriilidis; (II) Administrative support: All authors; (III) Provision of study materials or patients: P Gavriilidis, A Askari; (IV) Collection and assembly of data: P Gavriilidis; (V) Data analysis and interpretation: P Gavriilidis, A Askari; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*Correspondence to:* Paschalis Gavriilidis. Department of Hepato-Pancreato-Biliary and Transplant Surgery, Queen Elizabeth University Hospitals Birmingham NHS Foundation Trust, Birmingham, B15 1NU, UK. Email: pgavrielidis@yahoo.com.

**Abstract:** Cholangiocarcinoma (CC) is the second most common primary liver tumour. High-quality guidelines are essential for effective patient stratification and individualised treatment. This study aimed to appraise the methodological quality of existing guidelines for the resection of CC using the Appraisal of Guidelines for Research & Evaluation (AGREE II) instrument. A systematic search of the literature in Cochrane, PubMed, Google Scholar, and Embase was performed. Assessment of the clinical practice guidelines (CPGs) and consensuses was performed using the AGREE II instrument by four clinicians experienced in surgical practice and the AGREE II appraisal method. Literature searches identified 13 guidelines of highly variable quality according to the AGREE II criteria. The guidelines scored well in certain domains such as scope & purpose (median score across all guidelines; 65%), clarity of presentation (76%), and editorial independence (56%). However, they scored poorly for applicability (13%), rigour of development (30%), and stakeholder involvement (39%). None of the 13 guidelines was recommended universally for use without modification. Overall, the methodological quality of guidelines on the surgical management of CC is poor. Future updates should address and modify shortcomings detected by the AGREE II instrument, thereby facilitating better patient stratification and individualised treatment strategies.

**Keywords:** Cholangiocarcinoma (CC); clinical practice guidelines (CPGs); Appraisal of Guidelines for Research and Evaluation instrument (AGREE II instrument)

Submitted May 21, 2019. Accepted for publication Jul 26, 2019.
doi: 10.21037/hbsn.2019.09.06
**View this article at:** http://dx.doi.org/10.21037/hbsn.2019.09.06

## Introduction

Cholangiocarcinoma (CC) is a malignant tumour arising from epithelial lining of the biliary system that accounts 3% of all gastrointestinal tumours. According to its location, it can be divided into intrahepatic CC (ICC), which accounts for 20–25%, hilar CC (50–60%), and distal extrahepatic CC (20–25%) (1,2).

It is reported that the incidence rate of all forms of CC is demonstrating an increasing trend (3). However, coding misclassification of Klatskin tumour as ICC may have resulted in a skewed incidence rate by overestimating ICC by 13% and underestimating extrahepatic CC by 15%, (3,4). In 2011, the Institute of Medicine (US) revised the 21-year-old definition of clinical practice guidelines (CPGs) (Institute of Medicine, 1990) as follows, "Clinical practice guidelines are statements that include recommendations

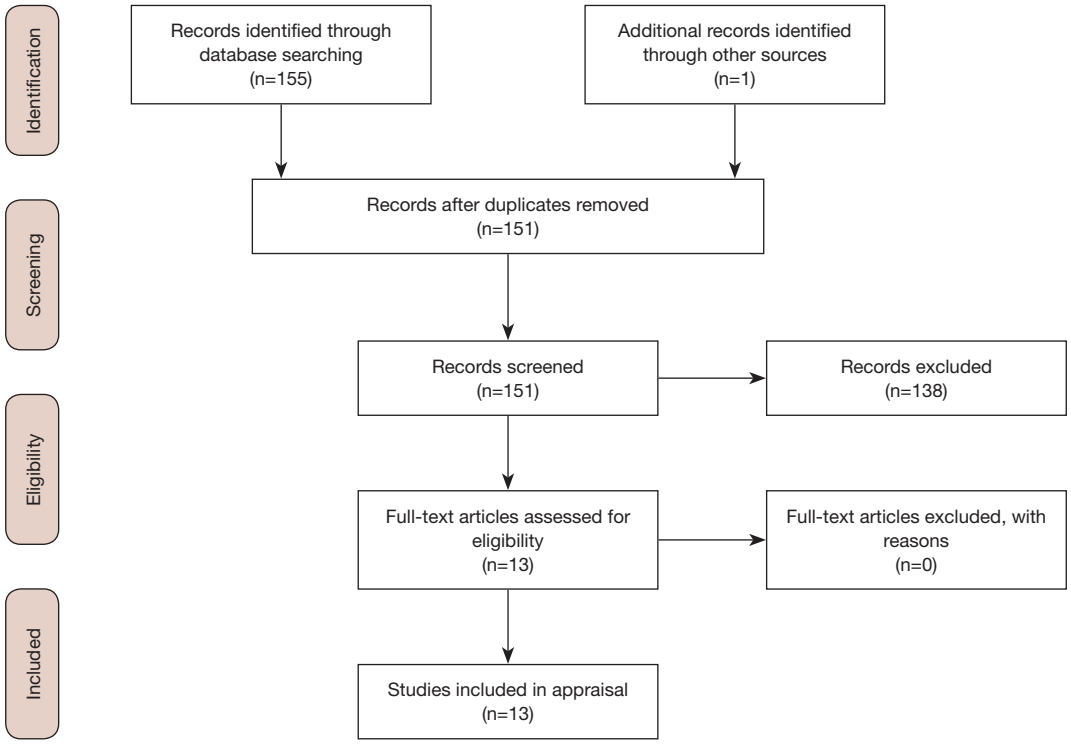


**Figure 1** Diagram of the search strategy.

intended to optimize patient care that are informed by a systematic review of the evidence and an assessment of the benefits and harms of alternative care options" (5).

The Appraisal of Guidelines for Research and Evaluation (AGREE) II instrument is the latest of more than 40 tools for the appraisal of CPGs (6,7). The AGREE instrument and its further refinements is the only CPG appraisal tool that has been developed and validated internationally, formally endorsed by several organizations including the WHO Advisory Committee on Health Research, and used by many guideline development groups (8,9). Detailed information is available on the AGREE web site (www.agreetrust.org).

The aim of the present study was to evaluate the quality of current CC guidelines, with a primary focus on resection, using the AGREE II instrument. The present study focused on the methodological analysis and did not analyse the recommended practices.

## Methods

### *Study selection and data review extraction*

A systematic review of the literature in Cochrane, PubMed, Embase, and Google Scholar (including studies of the last 20 years) was conducted to identify guidelines using Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria (10). References of retrieved articles were also searched manually for further guidelines. After independent evaluation of the CPGs by PG and RS, the following data were extracted: country of origin, year of publication, development and/or revision committee, evaluation measures, and funding sources. Evidenced-based CPGs in the English language pertaining to the resection of CC were included (*Figure 1*, *Tables 1* and *2*).

### *Appraisal of guidelines*

The AGREE II tool comprises 23 items divided into six domains: scope and purpose, stakeholder involvement, rigor of development, clarity of presentation, applicability, and editorial independence. For further details regarding the criteria used to describe and evaluate the 6 domains and the 23 consisted items, please see *Figure S1*: AGREE reporting checklist. After undergoing online training (www.agreetrust.org) to ensure appraisal standardisation, four appraisers (PG, AA, RS, KR), as recommended by the AGREE

 

**Table 1** Characteristics of the studies

| Guideline, year | Developers | Content | Guidelines review | Scoring system | Evidence based |
|---|---|---|---|---|---|
| Asian-Pacific Consensus, 2012 | Asian-Pacific consensus | Perihilar CC, E, D&T | A modified Delphi process | Levels of evidence | Systematic literature review, and consensus panel |
| Brazilian, 2015 | Brazilian Gastrointestinal Tumor Group | ICC, Hilar CC, D&T, FU | NR | Level of evidence, GRADE | Systematic literature review, expert panel |
| BSG, 2012 | BSG | CC, E, D&T, FU | BSG peer group review | Oxford + GRADE | Systematic literature review |
| Chinese Consensus, 2014 | A consensus of surgical specialists from China | CC, E, D&T | NR | NR | NR |
| EASL, 2014 | European association for the study of the liver | ICC, E, D&T | EASL reviewers | Oxford + GRADE | Systematic literature review and expert panel |
| ESMO, 2016 | ESMO Guidelines Committee | Biliary cancer, E, D&T, FU | NR | Levels of evidence, GRADE | ESMO-Guidelines-Methodology Expert panel |
| AHPBA Consensus for Hilar CC, 2014 | AHPBA Committee | Hilar CC, E, D&T | NR | NR | Expert panel |
| AHPBA Consensus for ICC, 2014 | AHPBA Committee | ICC, E, D&T | NR | NR | Expert panel |
| Italian Guidelines, 2010 | AIGO, AIOM, AIRO | CC, E, D + T | NR | NR | Expert panel |
| Japanese guidelines, 2008 | Japanese experts | Biliary cancer, E, D + T | NR | GRADE | Expert panel |
| NCCN, 2015 | National Comprehensive Cancer Network | CC, GBC, HCC | Members of NCCN | Oxford + GRADE | Systematic literature review, expert panel |
| Palliative Surgery for Hilar CC, 2008 | University of Mainz, Germany | CC, D&T | NR | NR | Expert panel |
| SEOM, 2015 | Socidad Española de Oncologia Medica | CC, GBC, E, D&T | NR | NR | Expert panel |

AHPBA, American Hepato-Pancreato-Biliary Association; SEOM, Socidad Española de Oncologia Medica; ESMO, European Society of Medical Oncology; EASL, European Association for the Study of the Liver; NCCN, National Comprehensive Cancer Network; AIGO, Italian Association of Hospital Gastroenterology; AIOM, Italian Association of Medical Oncology; AIRO, Association of Oncological Radiotherapy; BSG, British Society of Gastroenterology; CC, cholangiocarcinoma; ICC, intrahepatic cholangiocarcinoma; GBC, gallbladder cancer; E, epidemiology; D&T, diagnosis & treatment; FU, follow-up; NR, non-reported.

II consortium, evaluated the guidelines, independently using the AGREE II tool (September 2013 version). As per the AGREE II manual, discrepancies of more than 2 standard deviations (SDs) were resolved through dialogue. Authors had the ability to change their entry after group discussion. Domain scores were calculated with the following formula:

$$\frac{(\text{Obtained score} - \text{Minimum possible score})}{(\text{Maximum possible score} - \text{Minimum possible score})} \times 100\%$$

## Results

Search results yielded a total of 13 guidelines, eight of which were produced by multi-national organizations (*Figure 1*, *Table 1*) (11-23). Overall, the guidelines scored poorly: the median overall score was just 43%, with the highest overall score of 82% given to the British Society of Gastroenterology (BSG) guidelines, followed by 79% for the Asia-Pacific and European Association for the Study of the Liver (EASL) guidelines. Of the 13 guidelines, eight

 

**Table 2** Bismuth-Corlette and MSKCC classifications

| Classification | Definition |
|---|---|
| Modified Bismuth-Corlette Classification for Hilar Cholangiocarcinoma | |
| Type I | Below the confluence |
| Type II | Confined to confluence |
| Type IIIA | Extension into right hepatic duct |
| Type IIIB | Extension into left hepatic duct |
| Type IV | Extension into right and left hepatic ducts |
| Proposed modified T stage criteria for Hilar Cholangiocarcinoma by MSKCC | |
| T1 | Tumour confined to the confluence and/or right or left hepatic duct without portal vein involvement or atrophy |
| T2 | Tumour confined to the confluence and/or right or left hepatic duct with ipsilateral liver atrophy. No portal vein involvement demonstrated |
| T3 | Tumor confined to the confluence and/or right or left hepatic duct with ipsilateral portal vein branch involvement with/without associated ipsilateral lobar atrophy. No main portal vein involvement. (Occlusion, invasion, or encasement) |
| T4 | Any of the following: (I) tumor involving both right and left hepatic ducts up to secondary biliary radicals bilaterally; (II) main portal vein encasement |

MSKCC, Memorial Sloan Kettering Cancer Center.

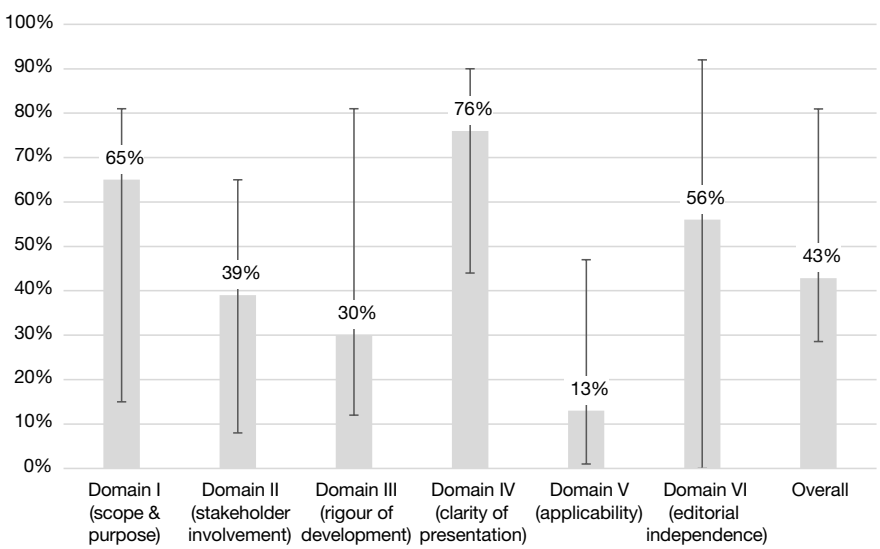


**Figure 2** Chart demonstrating median overall scores for each domain. Error bars denote the lowest and highest scoring guidelines.

scored under 50%. Median scores were particularly low in the following domains: II: stakeholder involvement (39%); III: rigor of development (30%); and V: applicability (13%). Domain IV: clarity of presentation and domain I: scope & purpose were the highest scoring at 76% and 65%, respectively (*Figure 2*).

### *Domain I: scope & purpose*

In this domain, questions pertain to the aims and objectives of the guidelines and the target users and population. Generally, the majority of the guidelines performed well, and the median score was 65%. The BSG scored the highest (81%), whilst the SEOM (Spanish Society of

 

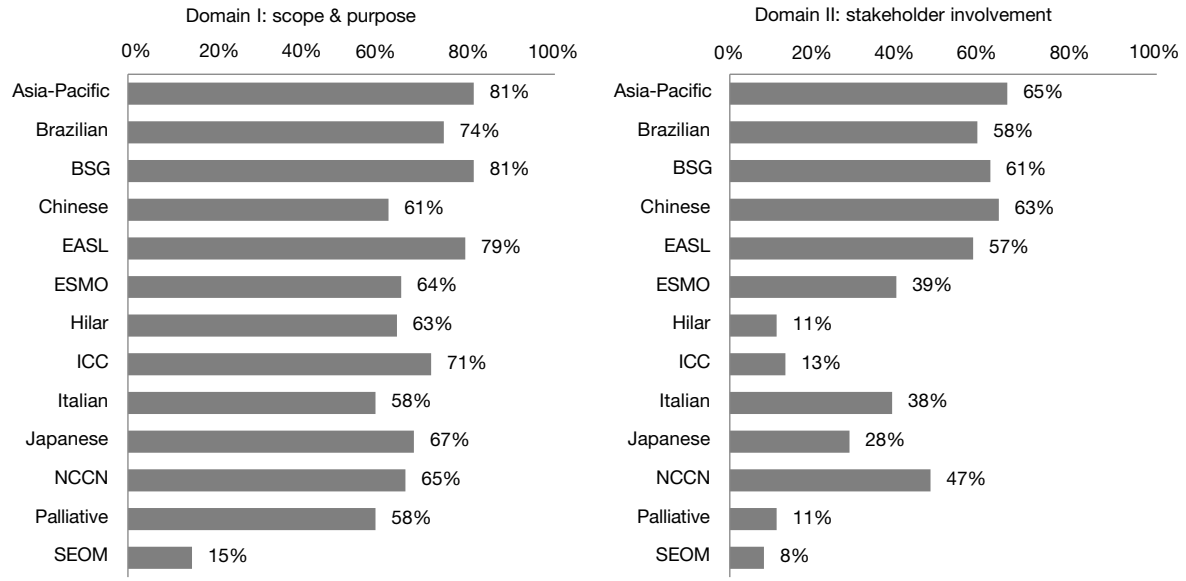


**Figure 3** Scores for domain I: scope & purpose and domain II: stakeholder involvement.

Medical Oncology) scored the lowest at 15% (*Figure 3*).

### *Domain II: stakeholder involvement*

An integral part of the AGREE II scoring checklist is the involvement of the relevant stakeholders in the guideline production process. Overall, the median score was 39% with the Asia-Pacific, Chinese, and BSG guidelines yielded the highest scores, at 65%, 63%, and 61%, respectively. The SEOM guidelines scored the lowest at 8% (*Figure 3*).

### *Domain III: rigor of development*

This was one of the lowest scoring domains. The overall score across all 13 guidelines was 30%, with the Asia-Pacific guidelines scoring the highest (81%) as they clearly laid out the methodology of the development of the guidelines from the evidence found in the literature (*Figure 4*). The Hilar guidelines had the lowest score at only 12%.

### *Domain IV: clarity of presentation*

Domain IV was the highest scoring domain with a median score of 76%. Among all of the guidelines, the Asia-Pacific and Japanese guidelines both scored 90%. The Hilar consensus statement received the lowest score of 44% (*Figure 4*).

### *Domain V: applicability*

Scores in this domain were the lowest of all, with the median score at just 13%. All of the guidelines scored less than 50%, with the highest scoring guideline being the Asia-Pacific guideline at 47% (*Figure 5*). The two lowest scoring guidelines scored a mere 1% (hilar and ICC-intrahepatic CC guidelines).

### *Domain VI: editorial independence*

Scores in domain VI were generally reasonable at 56%. The highest score (92%) was achieved by the European Society of Medical Oncology (ESMO) (*Figure 5*). One guideline scored 0% (guidelines for palliative surgery of CC).

### *Recommendation for use*

None of the 13 guidelines was recommended universally for use without modification. The Asia-Pacific and EASL guidelines scored a definitive 'Yes' by three appraisers and 'Yes, with modification' by the fourth appraiser. Five of the 13 guidelines (Hilar, ICC, Japanese, Palliative, and SEOM) scored a unanimous 'No' by the appraisers (*Figure 6*).

### *Score discrepancies*

Discrepancies in scoring across appraisers were low overall;

 

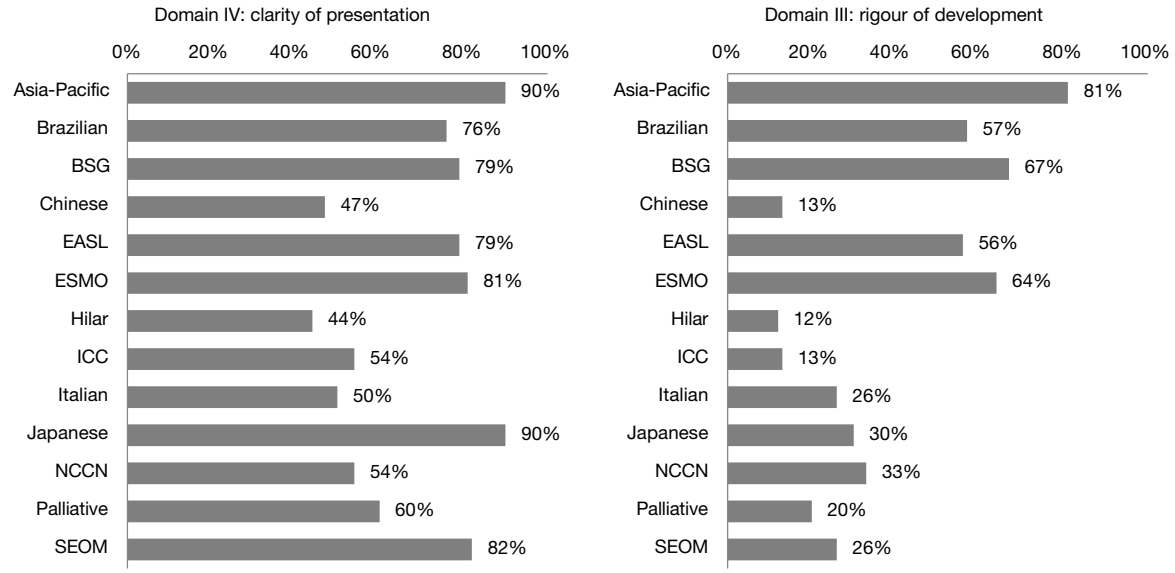


**Figure 4** Scores for domain III: rigor of development and domain IV: clarity of presentation.

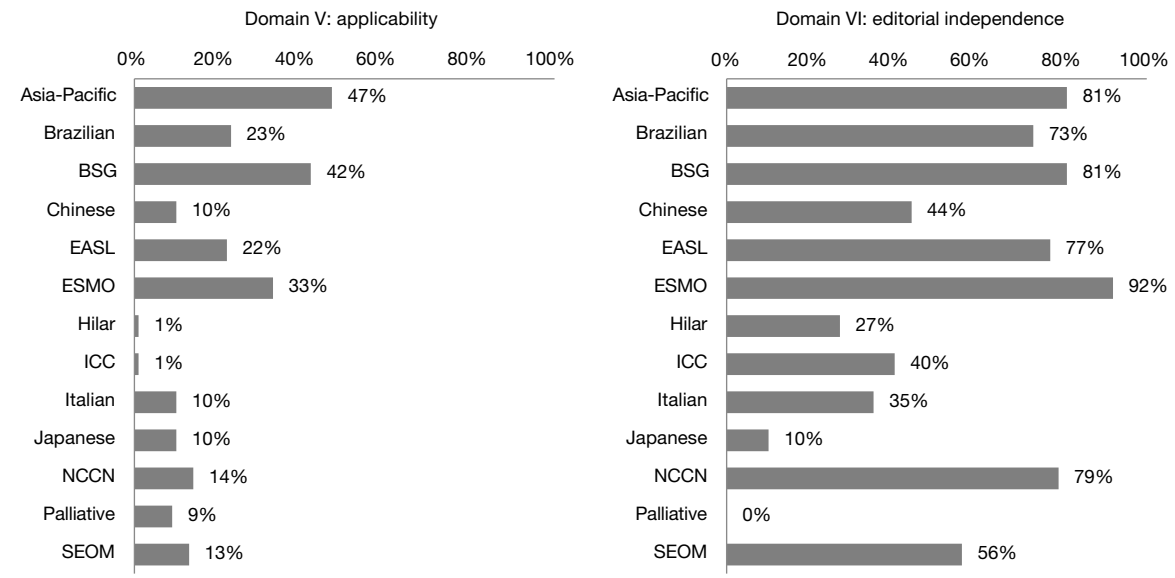


**Figure 5** Scores for domain V: applicability and domain VI: editorial independence.

therefore, no further rounds of re-scoring or discussion were required to resolve issues.

## Discussion

The present study is the first appraisal of the current CC guidelines using the validated AGREE II instrument. Overall, the quality of the guidelines as assessed by the AGREE II evaluation checklist was mediocre at best, with a median total score of only 43%. The BSG (82%), EASL (79%), and Asia-Pacific (79%) guidelines scored the highest overall score. Generally, the guidelines scored poorly in the domains of applicability, rigour of development, and stakeholder involvement, at 13%, 30%, and 39%, respectively. The highest scores were observed in the domains of clarity of presentation and scope and purpose, at

 

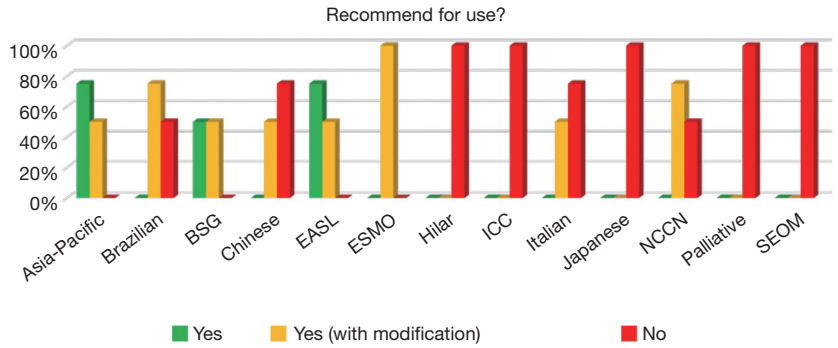


**Figure 6** Chart demonstrating recommendations for use for each guideline.

76% and 65%, respectively.

These findings highlight that careful attention and further developmental work on existing guidelines is required, particularly in the areas of clinical implementation and the involvement of patients/advocacy groups. Another area in which the guidelines are performing particularly poorly is their rigour of development. Often, the guidelines did not stipulate how they arrived at their recommendations. This is likely surprising given both the importance of transparency in guideline development and existence of validated systems for evaluating the scientific literature. One such system is the GRADE system, which has been specifically developed for the evaluation of evidence and classification of recommendations for guideline development (24). Nonetheless, only seven of 13 guidelines used the GRADE system to evaluate the quality of evidence and categorise the strength of recommendations, namely the Brazilian, NCCN, EASL, BSG, and Japanese guidelines (12,15,17,20,22).

Two others, the ESMO and the Asia-Pacific guidelines used alternative systems. The ESMO guidelines used the US Public Health Service grading system, and the Asia-Pacific guidelines categorised the evidence and classified the recommendations using a voting system based on a modification of the Canadian Task Force on the Periodic Health Examination (11,19). Four of the guidelines did not use any system at all (13,14,17,18). Consequently, the median score for rigour of development was only 38%. Guidelines that used a system to evaluate the evidence (i.e., the GRADE system) naturally scored higher, whilst those that did not employ any discernible system, such as the Hilar guidelines, scored poorly (12%).

Another area in which guidelines appear to perform universally poorly is stakeholder involvement, as has been extensively documented in the literature. The median score for this domain was just 39%, with the SEOM guidelines scoring only 8%. In the age of patient-centred care, patient autonomy, and informed consent, this lack of engagement with patients, patient advocacy groups, and the general public is concerning. In many instances, even other professionals involved in the treatment of patients with the condition were not consulted and the entire guideline was written by members of a single specialty. Although such guidelines can produce specialised recommendations from one aspect of care, i.e., surgical or oncological, they may miss other aspects not immediately within the purview of their specialty, which contradicts the ethos of holistic and multi-disciplinary care. Furthermore, very little guidance is provided by the guidelines. Of the 13 presented in this study, only two (Brazilian and ESMO) gave any recommendations for the follow-up and long-term management of these patients.

Another critical element of guidelines for the management of biliary tree pathology lies in the difference between benign and malignant disease and how they should be managed. For example, immunoglobulin G4 cholangiopathy is a multisystem inflammatory disorder that may present with intrahepatic biliary strictures in 51% of cases and proximal extrahepatic ducts in 49%; this disease should always be included in the differential diagnosis of biliary strictures (25). The BSG guidelines alone stressed the importance of differentiating between benign and malignant strictures. Such oversights highlight the limitations of the current guidelines in their scope and depth. In addition, the Japanese guidelines strongly recommend biopsy or cytology before surgery in order to differentiate malignant from benign strictures. Nakayama *et al.* reported that 10% of suspected and operable CCs were benign strictures (26). The remaining guidelines do not give recommendations regarding preoperative biopsy.

 

Moreover, staging laparoscopy is a useful tool to avoid unnecessary operations, though only the Asia-Pacific guidelines recommend the use of staging laparoscopy. In terms of ICC, the NCCN guidelines recommend colonoscopy and gastroscopy to rule out metastases from an asymptomatic gastrointestinal tumour. None of the other guidelines recommend it.

Most of the guidelines used the 7$^{th}$ AJCC staging system. It is reported that the main limitations of the AJCC are the definition of resectability and prediction of survival (27). The Blumgart staging system, which is based on the extent of biliary duct involvement by the tumour, the presence or absence of portal involvement, and the presence or absence of lobar atrophy, can define resectability and predict survival more accurately than the AJCC and Bismuth-Corlette staging systems (*Table 2*) (27). However, none of the guidelines proposed the above staging system.

The guidelines define the following as risk factors for CC: primary sclerosing cholangitis, parasitic infestations, hepatolithiasis, choledochal cysts, pancreatobiliary maljunction (PBM), toxins, and hepatitis B (HBV) and C (HCV) infections. However, only the BSG guidelines recommend the surveillance of patients with primary sclerosing cholangitis. In cases of PBM and choledochal cysts, the Japanese guidelines recommend cholecystectomy and extrahepatic common bile duct excision to prevent cancer development.

ICC and hepatocellular carcinoma (HCC) share many risk factors such as HBV, HCV, and cirrhosis. Guidelines on HCC recommend surveillance every six months with tumour markers and imaging modalities for high-risk patients (28). However, no such recommendation for surveillance of CC exists in any of the current guidelines.

Currently, there is no solid evidence to support standard lymph node dissection in patients with CC (29). However, NCCN guidelines suggest lymph node dissection to achieve better prognosis.

Expert consensus on ICC recommends that lymph node dissection be a standard part of the surgical management of patients with ICC.

It is reported that ≥7 lymph nodes are sufficient for the prognostic staging of hilar CC (30).

Another area in which guidelines scored particularly poorly was advice and guidance on how to implement the recommendations. This is an issue that has plagued many a guideline across a variety of medical sub-specialties, as well documented in the literature (28). Unfortunately, the CC guidelines are no exception. Very few guidelines documented the human and material resources required for the implementation of the recommendations or gave clear instructions as to how the recommendations could be put into action, leaving the readership with no real sense of direction or where to start.

There are even further shortcomings in the current CC guidelines. Perhaps most striking is the quality of the evidence on which many of the guidelines are based. Unfortunately, there is currently a distinct lack of randomised control trial data for many of the recommendations in place. A large proportion of the key points in many of the guidelines are based on observational data of potentially questionable reliability. If guidelines are to improve further, bold and rigorous studies within the boundaries of ethical consideration are required to further our understanding of the management of CC.

Our study has several limitations, some of which can be attributed to the very nature of the AGREE system. For example, it is debatable whether every domain in the AGREE system should carry the same weight in terms of scores as some of the others. The checklist has been criticised for its assumption that all domains are equally important for determining high-quality guidelines. Other criticisms of the checklist are that although assessors fill in their respective scores independently of each other, there is still the possibility of bias (positive or negative) and a certain level of subjectivity in the scoring. For example, certain guidelines from reputable international bodies may hold a favourable stance in the assessor's mind before the assessor even begins to undertake the scoring. Conversely, a guideline from a less well established body may not score as highly due to a lack of 'prestige'. A further limitation of our study is that we only selected guidelines in the English language for reasons of practicality; as such, we may have missed potentially high-quality guidelines presented in other languages.

Of note, The Asia-Pacific guidelines performs particularly well because their guideline is well set out, well written, thorough and easy to read. The evidence they base their recommendation on is no different to that found in many of the other guidelines, however what the Asia-Pacific guideline does well that others do not is the following:

(I) They clearly explain in detail the quality of the evidence they base their recommendations on, the degree of consensus and their methodology as to how they came about making the recommendations.
(II) They also rate their own recommendations in terms of quality and how strongly they recommend

 

a particular practice.

(III) The guideline is well presented, thorough in terms of addressing all the relevant aspects of investigation, treatment and outcome.

As a result of the above, the Asia-Pacific consistently scores well throughout most of the Domains.

Recently, Idrees *et al.* (31) evaluated the impact of the centralisation of care and compliance with NCCN guidelines for resected CCs on long-term survival. It was reported that over time, in the USA, compliance with NCCN guidelines increased. In particular, for the period of 2004–2007, the compliance was 30%, which increased to 46% in 2011–2013. Of note, five-year overall survival was 45% in the patients who received NCCN-compliant surgical management, as compared to 40% in those who did not receive surgical care according to the guidelines. Interestingly, it was reported that the centralisation of care contributed only 8% of the improvement in survival, while compliance with guidelines improved survival by 17% (31).

## Conclusions

The quality of the current guidelines for CC is generally poor or based on relatively low-quality evidence. It is imperative that future updated guidelines rely on high-quality trial data and take a multi-disciplinary approach by including patients and advocacy groups in the formulation of recommendations. Furthermore, a clear plan as to how to put the recommendations into practice (in both resource-rich and poor regions of the world) is desperately needed.

## Acknowledgments

*Funding:* None.

## Footnote

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at https://hbsn.amegroups.com/article/view/10.21037/hbsn.2019.09.06/coif). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.



## References

1. Bismuth Henri. Hepatobiliary malignancy. Edward Arnold. 1994.
2. de Groen PC, Gores GJ, LaRusso NF, et al. Biliary tract cancers. N Engl J Med 1999;341:1368-78.
3. Khan SA, Emadossadaty S, Ladep NG, et al. Rising trends in cholangiocarcinoma: Is the ICD classification system misleading us? J Hepatol 2012;56;848-54.
4. Welzel TM, McGlynn KA, Hsing AW, et al. Impact of classification of hilar cholangiocarcinomas (Klatskin tumors) on the incidence of intra- and extrahepatic cholangiocarcinoma in the United States. J Natl Cancer Inst 2006;98:873-75.
5. Graham R, Mancher M, Miller Wolman D, et al. Clinical practice guidelines we can trust. Institute of Medicine (US) Committee on Standards for developing trustworthy clinical practice guidelines;Washington (DC): The National Academies press (US); 2011.
6. Appraisal of Guidelines for Research and Evaluation II 2013. Available online: http//www.agreetrust.org
7. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting, and evaluation in health care. Prev Med 2010;51:421-4.
8. Coroneos CJ, Voineskos SH, Cornacchi SD, et al. Users' guide to the surgical literature: how to evaluate clinical practice guidelines. J Can Chir 2014;57:280-6.
9. 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. Qual Saf Health Care 2010;19:e58.
10. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009;151:264-9.
11. Valle JW, Borbath I, Khan SA, et al. Biliary cancer. ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2016;27:v28-37.
12. Riechelmann R, Coutinho A, Weschenfelder G, et al.

 

Guideline for the management of bile duct cancers by the Brazilian gastrointestinal tumor group. Arq Gastroenterol 2016;53:5-9.
13. Mansour JC, Aloia TA, Crane CH, et al. Hilar adenocarcinoma: Expert consensus statement. HPB 2015;17:691-9.
14. Weber SM, Ribero D, O' Reilly EM, et al. Intrahepatic Cholangiocarcinoma: expert consensus statement. HPB 2015;17:669-80.
15. Benson AB 3rd, Abrams TA, Ben-Josef E, et al. NCCN clinical practice guidelines in oncology. Hepatobiliary Cancers. J Natl Compr Canc Netw 2009;7:350-91.
16. Benavides M, Anton A, Gallego J, et al. Biliary tract cancers: SEOM clinical guideline. Clin Transl Oncol 2015;17:982-7.
17. Bridgewater J, Galle PR, Khan SA, et al. Guidelines for the diagnosis and management of intrahepatic cholangiocarcinoma. J Hepatol 2014;60:1268-89.
18. Cai JQ, Cai SW, Cong WM, et al. Diagnosis and treatment of cholangiocarcinoma: A consensus from surgical specialists of China. J Huazhong Univ Sci technol 2014;34:469-75.
19. Rerknimitr R, Angsuwatcharakon P, Ratanachu-ek T, et al. Asian-Pacific consensus recommendations for endoscopic and interventional management of hilar cholangiocarcinoma. J Gastroenterol Hepatol 2013;28:593-607.
20. Khan SA, Davidson BR, Golden RD, et al. Guidelines for the diagnosis and treatment of cholangiocarcinoma: An update. Gut 2012;61:1657-69.
21. Alvaro D, Cannizaro R, Labianca R, et al. Cholangiocarcinoma: A position paper for the Italian society of gastroenterology (SIGE), the Italian Association of Hospiatal Gastroenterology (AIGO), the Italian Association of Medical Oncology (AIOM) and the Italian Association of Oncological Radiotherapy (AIRO). Dig Liver Dis 2010;831-8.
22. Kondo S, Takada T, Miyazaki M, et al. Guidelines for the management of biliary tract and ampullary carcinoma. Surgical treatment. J Hepatobiliary Pancreat Surg 2008;15:41-54.
23. Witzigmann H, Lang H, Lauer H. Guidelines for palliative surgery of cholangiocarcinoma. HPB 2008;10:154-60.
24. Norris SL, Bero L. GRADE Methods for Guideline Development: Time to Evolve? Ann Intern Med 2016;165:810-1.
25. Ghazale A, Chari ST, Zhang L, et al. Immunoglobulin G4-associated cholangitis: clinical profile and response to therapy. Gastroenterology 2008;134:706-15.
26. Nakayama A, Imamura H, Shimada R, et al. Proximal bile duct stricture disguised as malignant neoplasm. Surgery 1999;125:514-21.
27. Jarnagin WR, Fong Y, DeMatteo RP, et al. Staging, resectability, and outcome in 225 patients with hilar cholangiocarcinoma. Ann Surg 2001;234:507-17; discussion 517-9.
28. 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;67:991-8.
29. Amini N, Ejaz A, Spolverato G, et al. Management of the lymph nodes during resection of hepatocellular carcinoma and intrahepatic cholangiocarcinoma: A systematic review. J Gastrointest Surg 2014;18:2136-48.
30. Kambakamba P, Linecker M, Slankamenac K, et al. Lymph node dissection in resectable perihilar cholangiocarcinoma: A systematic review. Am J Surg 2015;210:694-701.
31. Idrees JJ, Merath K, Gani F, et al. Trends in centralization of surgical care and compliance with National Cancer Center Network guidelines for resected cholangiocarcinoma. HPB 2019;21:981-9.

**Cite this article as:** 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(2):126-135. doi: 10.21037/hbsn.2019.09.06

 

Supplementary

# AGREE Reporting Checklist
2016

*This checklist is intended to guide the reporting of clinical practice guidelines.*

| CHECKLIST ITEM AND DESCRIPTION | REPORTING CRITERIA | Page # |
|---|---|---|
| ***DOMAIN 1: SCOPE AND PURPOSE*** | | |
| **1. OBJECTIVES**<br>*Report the overall objective(s) of the guideline. The expected health benefits from the guideline are to be specific to the clinical problem or health topic.* | ☐ Health intent(s) (i.e., prevention, screening, diagnosis, treatment, etc.)<br>☐ Expected benefit(s) or outcome(s)<br>☐ Target(s) (e.g., patient population, society) | |
| **2. QUESTIONS**<br>*Report the health question(s) covered by the guideline, particularly for the key recommendations.* | ☐ Target population<br>☐ Intervention(s) or exposure(s)<br>☐ Comparisons (if appropriate)<br>☐ Outcome(s)<br>☐ Health care setting or context | |
| **3. POPULATION**<br>*Describe the population (i.e., patients, public, etc.) to whom the guideline is meant to apply.* | ☐ Target population, sex and age<br>☐ Clinical condition (if relevant)<br>☐ Severity/stage of disease (if relevant)<br>☐ Comorbidities (if relevant)<br>☐ Excluded populations (if relevant) | |
| ***DOMAIN 2: STAKEHOLDER INVOLVEMENT*** | | |
| **4. GROUP MEMBERSHIP**<br>*Report all individuals who were involved in the development process. This may include members of the steering group, the research team involved in selecting and reviewing/rating the evidence and individuals involved in formulating the final recommendations.* | ☐ Name of participant<br>☐ Discipline/content expertise (e.g., neurosurgeon, methodologist)<br>☐ Institution (e.g., St. Peter's hospital)<br>☐ Geographical location (e.g., Seattle, WA)<br>☐ A description of the member's role in the guideline development group | |
| **5. TARGET POPULATION PREFERENCES AND VIEWS**<br>*Report how the views and preferences of the target population were sought/considered and what the resulting outcomes were.* | ☐ Statement of type of strategy used to capture patients'/publics' views and preferences (e.g., participation in the guideline development group, literature review of values and preferences)<br>☐ Methods by which preferences and views were sought (e.g., evidence from literature, surveys, focus groups)<br>☐ Outcomes/information gathered on patient/public information<br>☐ How the information gathered was used to inform the guideline development process and/or formation of the recommendations | |
| **6. TARGET USERS**<br>*Report the target (or intended) users of the guideline.* | ☐ The intended guideline audience (e.g. specialists, family physicians, patients, clinical or institutional leaders/administrators)<br>☐ How the guideline may be used by its target audience (e.g., to inform clinical decisions, to inform policy, to inform standards of care) | |
| ***DOMAIN 3: RIGOUR OF DEVELOPMENT*** | | |
| **7. SEARCH METHODS**<br>*Report details of the strategy used to search for evidence.* | ☐ Named electronic database(s) or evidence source(s) where the search was performed (e.g., MEDLINE, EMBASE, PsychINFO, CINAHL)<br>☐ Time periods searched (e.g., January 1, 2004 to March 31, 2008)<br>☐ Search terms used (e.g., text words, indexing terms, subheadings)<br>☐ Full search strategy included (e.g., possibly located in appendix) | |
| **8. EVIDENCE SELECTION CRITERIA**<br>*Report the criteria used to select (i.e., include and exclude) the evidence. Provide rationale, where appropriate.* | ☐ Target population (patient, public, etc.) characteristics<br>☐ Study design<br>☐ Comparisons (if relevant)<br>☐ Outcomes<br>☐ Language (if relevant)<br>☐ Context (if relevant) | |
| **9. STRENGTHS & LIMITATIONS OF THE EVIDENCE**<br>*Describe the strengths and limitations of the evidence. Consider from the perspective of the individual studies and the body of evidence aggregated across all the studies. Tools exist that can facilitate the reporting of this concept.* | ☐ Study design(s) included in body of evidence<br>☐ Study methodology limitations (sampling, blinding, allocation concealment, analytical methods)<br>☐ Appropriateness/relevance of primary and secondary outcomes considered<br>☐ Consistency of results across studies<br>☐ Direction of results across studies<br>☐ Magnitude of benefit versus magnitude of harm<br>☐ Applicability to practice context | |
| **10. FORMULATION OF RECOMMENDATIONS**<br>*Describe the methods used to formulate the recommendations and how final decisions were reached. Specify any areas of disagreement and the methods used to resolve them.* | ☐ Recommendation development process (e.g., steps used in modified Delphi technique, voting procedures that were considered)<br>☐ Outcomes of the recommendation development process (e.g., extent to which consensus was reached using modified Delphi technique, outcome of voting procedures)<br>☐ How the process influenced the recommendations (e.g., results of Delphi technique influence final recommendation, alignment with recommendations and the final vote) | |
| **11. CONSIDERATION OF BENEFITS AND HARMS**<br>*Report the health benefits, side effects, and risks that were considered when formulating the recommendations.* | ☐ Supporting data and report of benefits<br>☐ Supporting data and report of harms/side effects/risks<br>☐ Reporting of the balance/trade-off between benefits and harms/side effects/risks<br>☐ Recommendations reflect considerations of both benefits and harms/side effects/risks | |
| **12. LINK BETWEEN RECOMMENDATIONS AND EVIDENCE**<br>*Describe the explicit link between the recommendations and the evidence on which they are based.* | ☐ How the guideline development group linked and used the evidence to inform recommendations<br>☐ Link between each recommendation and key evidence (text description and/or reference list)<br>☐ Link between recommendations and evidence summaries and/or evidence tables in the results section of the guideline | |
| **13. EXTERNAL REVIEW**<br>*Report the methodology used to conduct the external review.* | ☐ Purpose and intent of the external review (e.g., to improve quality, gather feedback on draft recommendations, assess applicability and feasibility, disseminate evidence)<br>☐ Methods taken to undertake the external review (e.g., rating scale, open-ended questions)<br>☐ Description of the external reviewers (e.g., number, type of reviewers, affiliations)<br>☐ Outcomes/information gathered from the external review (e.g., summary of key findings)<br>☐ How the information gathered was used to inform the guideline development process and/or formation of the recommendations (e.g., guideline panel considered results of review in forming final recommendations) | |
| **14. UPDATING PROCEDURE**<br>*Describe the procedure for updating the guideline.* | ☐ A statement that the guideline will be updated<br>☐ Explicit time interval or explicit criteria to guide decisions about when an update will occur<br>☐ Methodology for the updating procedure | |
| ***DOMAIN 4: CLARITY OF PRESENTATION*** | | |
| **15. SPECIFIC AND UNAMBIGUOUS RECOMMENDATIONS**<br>*Describe which options are appropriate in which situations and in which population groups, as informed by the body of evidence.* | ☐ A statement of the recommended action<br>☐ Intent or purpose of the recommended action (e.g., to improve quality of life, to decrease side effects)<br>☐ Relevant population (e.g., patients, public)<br>☐ Caveats or qualifying statements, if relevant (e.g., patients or conditions for whom the recommendations would not apply)<br>☐ If there is uncertainty about the best care option(s), the uncertainty should be stated in the guideline | |
| **16. MANAGEMENT OPTIONS**<br>*Describe the different options for managing the condition or health issue.* | ☐ Description of management options<br>☐ Population or clinical situation most appropriate to each option | |
| **17. IDENTIFIABLE KEY RECOMMENDATIONS**<br>*Present the key recommendations so that they are easy to identify.* | ☐ Recommendations in a summarized box, typed in bold, underlined, or presented as flow charts or algorithms<br>☐ Specific recommendations grouped together in one section | |
| ***DOMAIN 5: APPLICABILITY*** | | |
| **18. FACILITATORS AND BARRIERS TO APPLICATION**<br>*Describe the facilitators and barriers to the guideline's application.* | ☐ Types of facilitators and barriers that were considered<br>☐ Methods by which information regarding the facilitators and barriers to implementing recommendations were sought (e.g., feedback from key stakeholders, pilot testing of guidelines before widespread implementation)<br>☐ Information/description of the types of facilitators and barriers that emerged from the inquiry (e.g., practitioners have the skills to deliver the recommended care, sufficient equipment is not available to ensure all eligible members of the population receive mammography)<br>☐ How the information influenced the guideline development process and/or formation of the recommendations | |
| **19. IMPLEMENTATION ADVICE/TOOLS**<br>*Provide advice and/or tools on how the recommendations can be applied in practice.* | ☐ Additional materials to support the implementation of the guideline in practice. For example:<br>◻ Guideline summary documents<br>◻ Links to check lists, algorithms<br>◻ Links to how-to manuals<br>◻ Solutions linked to barrier analysis (see Item 18)<br>◻ Tools to capitalize on guideline facilitators (see Item 18)<br>◻ Outcome of pilot test and lessons learned | |
| **20. RESOURCE IMPLICATIONS**<br>*Describe any potential resource implications of applying the recommendations.* | ☐ Types of cost information that were considered (e.g., economic evaluations, drug acquisition costs)<br>☐ Methods by which the cost information was sought (e.g., a health economist was part of the guideline development panel, use of health technology assessments for specific drugs, etc.)<br>☐ Information/description of the cost information that emerged from the inquiry (e.g., specific drug acquisition costs per treatment course)<br>☐ How the information gathered was used to inform the guideline development process and/or formation of the recommendations | |
| **21. MONITORING/ AUDITING CRITERIA**<br>*Provide monitoring and/or auditing criteria to measure the application of guideline recommendations.* | ☐ Criteria to assess guideline implementation or adherence to recommendations<br>☐ Criteria for assessing impact of implementing the recommendations<br>☐ Advice on the frequency and interval of measurement<br>☐ Operational definitions of how the criteria should be measured | |
| ***DOMAIN 6: EDITORIAL INDEPENDENCE*** | | |
| **22. FUNDING BODY**<br>*Report the funding body's influence on the content of the guideline.* | ☐ The name of the funding body or source of funding (or explicit statement of no funding)<br>☐ A statement that the funding body did not influence the content of the guideline | |
| **23. COMPETING INTERESTS**<br>*Provide an explicit statement that all group members have declared whether they have any competing interests.* | ☐ Types of competing interests considered<br>☐ Methods by which potential competing interests were sought<br>☐ A description of the competing interests<br>☐ How the competing interests influenced the guideline process and development of recommendations | |

From:
Brouwers MC, Kerkvliet K, Spithoff K, on behalf of the AGREE Next Steps Consortium. The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines. *BMJ* 2016;352:i1152. doi: 10.1136/bmj.i1152.

For more information about the AGREE Reporting Checklist, please visit the AGREE Enterprise website at www.agreetrust.org.

**Figure S1** AGREE reporting checklist 2016.