Urologia Internationalis

Original Paper

Urol Int 2021;105:31–40
DOI: 10.1159/000509431

Received: March 7, 2020
Accepted: June 7, 2020
Published online: August 21, 2020

Clinical Practice Guidelines for Bladder Cancer: A Systematic Review and Meta-Analysis Using the AGREE II Instrument

Philipp Maisch Margitta Retz Jürgen E. Gschwend Florestan Koll
Sebastian C. Schmid

Department of Urology, Rechts der Isar Medical Center, Technical University of Munich, Munich, Germany

Keywords
Urinary bladder neoplasms · Bladder cancer · Clinical practice guideline · AGREE II

Abstract
Context: Numerous health care organizations have established guidelines on diagnosis and treatment of bladder cancer. However, the lack of a standardized guideline development approach results in considerable differences of the guidelines' methodological quality. ***Objective:*** To assess the methodological quality of all relevant clinical practice guidelines (CPGs) for urinary bladder cancer and provide a reference for clinicians in choosing guidelines of high methodological quality. ***Evidence Acquisition:*** A systematic literature search was conducted in Medline via PubMed, 4 CPG databases, and 7 databases of interdisciplinary organizations. CPGs for non-muscle-invasive bladder cancer (NMIBC) and muscle-invasive bladder cancer (MIBC) with the topics screening, pathology, diagnosis, treatment, and aftercare published in English language between 2012 and 2018 were included. The CPG quality was analyzed using the Appraisal of Guidelines for Research and Evaluation (AGREE) II instrument. ***Evidence Synthesis:*** A total of 16 CPGs were included for the quality appraisal. Because of predefined criteria, 5 CPGs were "strongly recommended" (American Urological Association NMIBC, European Association of Urology [EAU] NMIBC, EAU MIBC, National Institute for Health and Care Excellence, and National Comprehensive Cancer Network), 4 CPGs were "weakly recommended" and 7 CPGs were "not recommended." ***Conclusions:*** The methodological quality of bladder cancer guidelines is diverse. Considering the rapid development of new therapies (e.g., immune checkpoint inhibitors), "living guidelines" of high methodological quality, such as the EAU NMIBC or MIBC guideline, will become more relevant in the future guideline's landscape.

Introduction

Bladder cancer (BC) is the ninth most common malignancy in the world. The incidence rate is 5.3 per 100,000 persons/years leading to an estimated 165,000 deaths per year [1–3]. Diagnosis and treatment of BC is complex and expensive, resulting in a significant burden for patients and health care systems. Therefore, it is crucial to optimize clinical decision-making, which can be supported

karger@karger.com
www.karger.com/uin

Philipp Maisch
Department of Urology, Rechts der Isar Medical Center
Technical University of Munich, Ismaninger Strasse 22
DE–81675 Munich (Germany)
philipp.maisch@mri.tum.de

Table 1. Content of the included guidelines for NMIBC or MIBC or any special topic

Guideline, year	Cancer entity			Topics included				
	NMIBC	MIBC	other topics	screening	pathology	diagnosis	treatment	aftercare
ACR pretreat [12]		⊗	ST					
ACR posttreat [13]	⊗	⊗	ST					⊗
Albertra NMIBC [14]	⊗		BCG				⊗	
Albertra MIBC [15]		⊗				⊗	⊗	⊗
AUA NMIBC [16]	⊗				⊗	⊗	⊗	⊗
AUA MIBC [17]		⊗					⊗	⊗
CUA (2015) [15]	⊗						⊗	
EAU NMIBC (2018) [16]	⊗				⊗	⊗	⊗	⊗
EAU MIBC (2018) [17]		⊗			⊗	⊗	⊗	⊗
EAU laser (2014) [18]	⊗		LT				⊗	
EAU robotic (2014) [19]		⊗	RARC				⊗	
ESMO (2014) [20]	⊗	⊗			⊗		⊗	⊗
NICE (2015) [21]	⊗	⊗		⊗		⊗	⊗	⊗
NCCN (2018) [22]	⊗	⊗				⊗	⊗	⊗
MDT (2013) [23]	⊗	⊗				⊗	⊗	
SEOM (2016) [24]		⊗				⊗	⊗	⊗

NMIBC, non-muscle-invasive bladder cancer; MIBC, muscle-invasive bladder cancer; ST, staging; RARC, robot-assisted radical cystectomy; LT, laser technologies; BCG, *Bacillus* Calmette-Guerin; ACR, American College of Radiology; AUA, American Urological Association; CUA, Canadian Urological Association; EAU, European Association of Urology; ESMO, European Society of Medical Oncology; NICE, National Institute for Health and Clinical Excellence; NCCN, National Comprehensive Cancer Network; MDT, multidisciplinary team; SEOM, Spanish Sociedad Española de Oncología de Médica.

by clinical practice guidelines (CPGs) as a helpful device. The Institute of Medicine defined CPGs as "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances" [4]. Their aim is to provide a well-considered, evidence-based, and unbiased summary of the current knowledge that is easily accessible as well as transparent for health professionals and patients in daily life [5]. Thus, the implementation of good guidelines can improve health care quality on the basis of up-to-date knowledge and reduce unnecessary interventions and costs.

In the last decade, notable human and financial resources have been raised by health care organizations and governments resulting in multiple, individual CPGs and consensus recommendations on identical tumor entities [6, 7]. Unfortunately, because of missing developing standards and quality characteristics, these CPGs differ in their methodological quality, leading to divergent recommendations. For common guideline users, it is nearly impossible to identify CPGs of good methodological quality in their everyday clinical routine. Thus, the methodological quality of guidelines should be reviewed critically before adopting recommendations into clinical routine. This process may help clinical decision-makers in the right choice of guidelines.

In order to assess the guidelines' methodological quality, the Appraisal of Guidelines for Research and Evaluation (AGREE) II instrument is an adequate tool [5, 8–10]. The AGREE II instrument is a paper-based checklist. First published in 2003 by a group of international guideline developers and researchers, it was updated in 2010 to improve its reliability and validity. AGREE II has been cited in over 600 publications and was endorsed by several health care organizations [11]. The methodological evaluation with AGREE II is divided into 6 domains including *scope and purpose*, *stakeholder involvement*, *rigor of development*, *clarity of presentation*, *applicability*, and *editorial independence* (Table 2). As a result of this analysis, CPGs are "strongly," "weakly," or "not" recommended. However, AGREE II is not able to validate the content, clinical quality, or applicability of CPGs. Thus, high methodological quality of a CPG does not necessarily correlate with clinical relevance, although it is likely that a high methodological quality leads to higher quality of the overall guideline.

To date, the published CPGs on bladder cancer differ significantly in their included content on tumor entities (Table 1) [12–27]. Most CPGs do not include all stages of

bladder cancer from non-muscle-invasive to metastatic disease or topics such as screening, pathology, diagnosis, treatment, and aftercare. Few CPGs only focus on special aspects of bladder cancer such as bladder instillation therapy or laser technology.

To our knowledge, no systematic quality assessment of CPGs for bladder cancer has been published so far. The aim of our study was to identify all relevant guidelines for bladder cancer and to perform a quality assessment using the AGREE II instrument. A meta-analysis of the results was conducted in order to identify bladder cancer guidelines of high methodological quality.

Evidence Acquisition

Inclusion Critieria

CPGs for non-muscle-invasive cancer (NMIBC) and muscle-invasive bladder cancer (MIBC) with the topics screening, pathology, diagnosis, treatment, and aftercare published in English language between 2012 and 2018 were included.

Exclusion Criteria

CPGs for nursing and upper urinary tract carcinoma, position statements, guideline summaries, and expert opinions such as publications from consensus development conferences were not evaluated.

Databases and Search Strategies

A systematic literature search was conducted from 2012 up to October 2017 in (a) Medline via PubMed; (b) 4 CPG databases: Guidelines International Network (GIN), National Guideline Clearinghouse (NGC), German Consortium of Scientific Medical Associations (AWMF), and NHS Evidence; and (c) 7 databases of interdisciplinary/expert organizations: European Association of Urology (EAU), Scottish Intercollegiate Guidelines Network (SIGN), National Institute for Health and Clinical Excellence (NICE), National Health and Medical Research Council (NHMRC), New Zealand Guidelines Group (NZGG), Cancer Council Australia (CCA), and Cancer Care Ontario (CCO).

An update literature search was performed in October 2018 of the guidelines found in the initial systematic search. Three CPGs (EAU NMIBC, EAU MIBC, and National Comprehensive Cancer Network [NCCN]) were available in an updated version.

The terms of search and the search strategy were adapted to the databases used. Terms of search including

Table 2. Explanation of the AGREE II domains and its 23 key items [9]

Domain 1. Scope and purpose
1. The overall objective(s) of the guideline is (are) specifically described
2. The health question(s) covered by the guideline is (are) specifically described
3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described
Domain 2. Stakeholder involvement
4. The guideline development group includes individuals from all the relevant professional groups
5. The views and preferences of the target population (patients, public, etc.) have been sought
6. The target users of the guideline are clearly defined
Domain 3. Rigor of development
7. Systematic methods were used to search for evidence
8. The criteria for selecting the evidence are clearly described
9. The strengths and limitations of the body of evidence are clearly described
10. The methods for formulating the recommendations are clearly described
11. The health benefits, side effects, and risks have been considered in formulating the recommendations
12. There is an explicit link between the recommendations and the supporting evidence
13. The guideline has been externally reviewed by experts prior to its publication
14. A procedure for updating the guideline is provided
Domain 4. Clarity of presentation
15. The recommendations are specific and unambiguous
16. The different options for management of the condition or health issue are clearly presented
17. Key recommendations are easily identifiable
Domain 5. Applicability
18. The guideline describes facilitators and barriers to its application
19. The guideline provides advice and/or tools on how the recommendations can be put into practice
20. The potential resource implications of applying the recommendations have been considered
21. The guideline presents monitoring and/or auditing criteria Domain 6. Editorial independence
22. The views of the funding body have not influenced the content of the guideline
23. Competing interests of guideline development group members have been recorded and addressed

AGREE II, Appraisal of Guidelines for Research and Evaluation II.

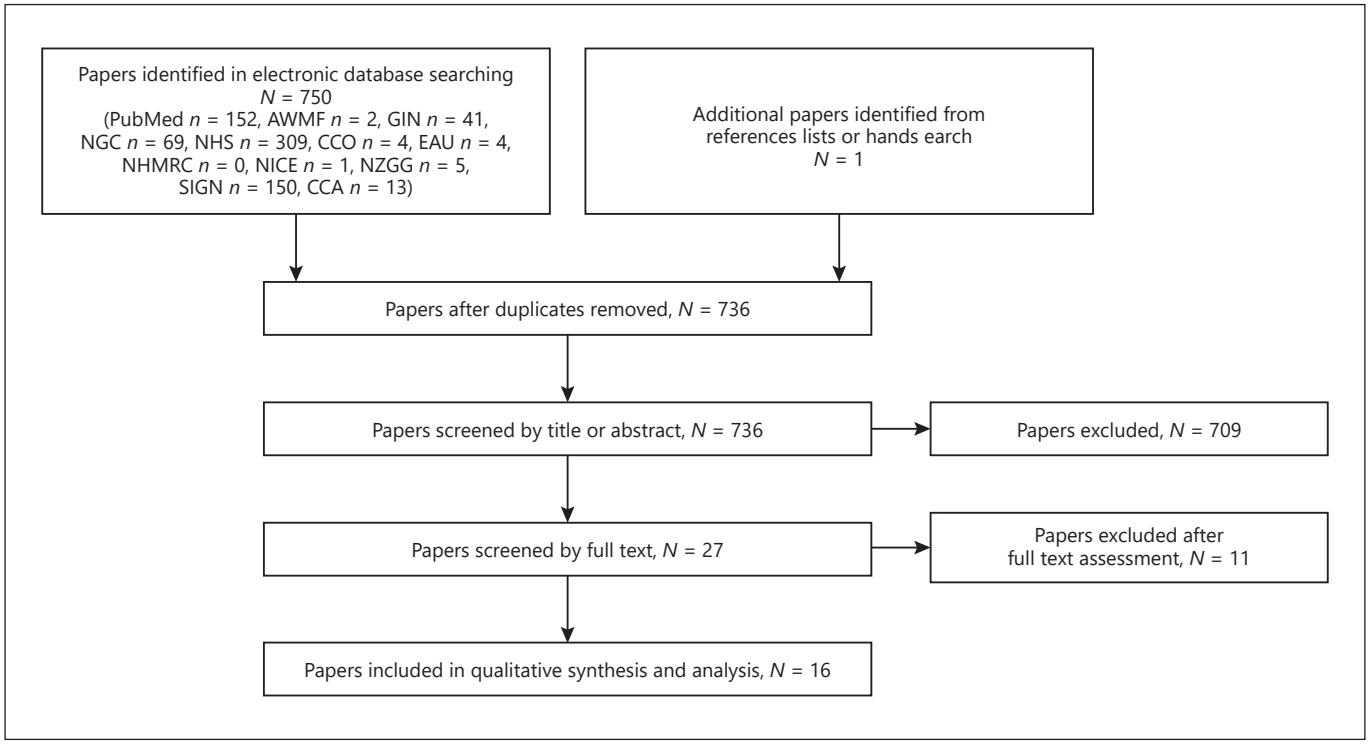


Fig. 1. PRISMA flowchart of the selection process for the publications identified by systematic literature search. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

MeSH terms were urinary bladder neoplasms, bladder cancer, guideline, and practice guideline. The complete search strategy for Medline via PubMed is attached in online suppl. 1; for all online suppl. material, see www.karger.com/doi/10.1159/000509431.

Literature Selection

According to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement, 2 independent reviewers (P.M. and S.C.S.) screened all publications found and applied predefined inclusion and exclusion criteria (Fig. 1) [28]. In case of discrepancies, the decision for inclusion was made by involvement of a third reviewer (MR).

Quality Assessment

The AGREE II instrument is a validated tool for the quality assessment of CPGs [5, 9, 10]. Within 6 domains, 23 key items are organized (Table 2). The 6 domains are *scope and purpose* (3 items), *stakeholder involvement* (3 items), *rigor of development* (8 items), *clarity of presentation* (3 items), *applicability* (4 items), and *editorial independence* (2 items).

Each of these 23 key items was assessed for every guideline on a 7-score rating scale (1 – strongly disagree; 7 – strongly agree) independently by 2 authors (P.M. and S.C.S.). Both authors were familiar with the use of the AGREE II instrument. If items were rated with a difference of >2 points, a consensus was made within a discussion for every author's rating. The evaluation of this assessment was performed by 1 author (MR).

Following the AGREE II regulations, the author's rating scores are expressed on percent scale from 0 to 100%. Calculations were performed with the 7-point AGREE II score calculator, available online [29].

A domain score of ≥50% was defined as the lowest threshold to rate the overall quality of a CPG as high. A CPG was "strongly recommended" if 4 or more domains scored ≥50%. With a minimum of 3 domains rated ≥50%, a guideline was "weakly recommended." CPGs were "not recommended" if they scored less. Since the AGREE II instrument does not set any benchmarks for the overall rating, this rating was based on an individual decision as recommended by AGREE II. The rationale for the ≥50% threshold was the use of this percentage in other publications and guideline development programs. Compared

Table 3. Results of the clinical practice guidelines evaluated with the AGREE II instrument

AGREE II domains/guidelines	Domain score in %						
	scope and purpose	stakeholder involvement	rigor of development	clarity of presentation	applicability	editorial independence	overall recommendation
ACR pretreat (2012)	17	25	14	56	0	50	Not
ACR posttreat (2014)	17	25	14	56	0	50	Not
Albertra NMIBC (2013)	97	11	48	86	13	58	Weakly
Albertra MIBC (2013)	97	11	47	92	10	63	Weakly
AUA NMIBC (2016)	86	56	49	83	10	58	Strongly
AUA MIBC (2017)	61	39	49	94	19	63	Weakly
CUA (2015)	42	39	24	56	10	0	Not
EAU NMIBC (2018)	64	42	52	97	13	63	Strongly
EAU MIBC (2018)	61	39	51	94	19	63	Strongly
EAU laser (2014)	39	58	16	53	2	58	Weakly
EAU robotic (2014)	31	6	16	53	17	58	Not
ESMO (2014)	8	6	0	33	0	21	Not
NICE (2015)	100	89	76	100	44	92	Strongly
NCCN (2018)	53	69	35	67	13	100	Strongly
MDT (2013)	67	14	9	39	0	38	Not
SEOM (2016)	39	6	7	47	0	13	Not
Median	57.0	32.0	29.5	61.5	10.0	58.0	

ACR, American College of Radiology; AUA, American Urological Association; CUA, Canadian Urological Association; EAU, European Association of Urology; ESMO, European Society of Medical Oncology; NICE, National Institute for Health and Clinical Excellence; NCCN, National Comprehensive Cancer Network; MDT, multidisciplinary team; SEOM, Spanish Sociedad Española de Oncología de Médica.

with the range of the median scores of the AGREE II domains (10.0–61.5%), the ≥50% threshold represents in our opinion an appropriate cutoff value for "strongly recommended" guidelines.

Statistical Analysis

For each AGREE II domain, the median score was calculated. The reason to analyze median scores of different domains in a meta-analysis was to compare the strength and weakness of different aspects of the presented guidelines. To estimate the interrater reliability, the average intraclass correlation including the 95% confidence interval was applied. Descriptive statistics were performed using SPSS, version 23 for Windows.

Results

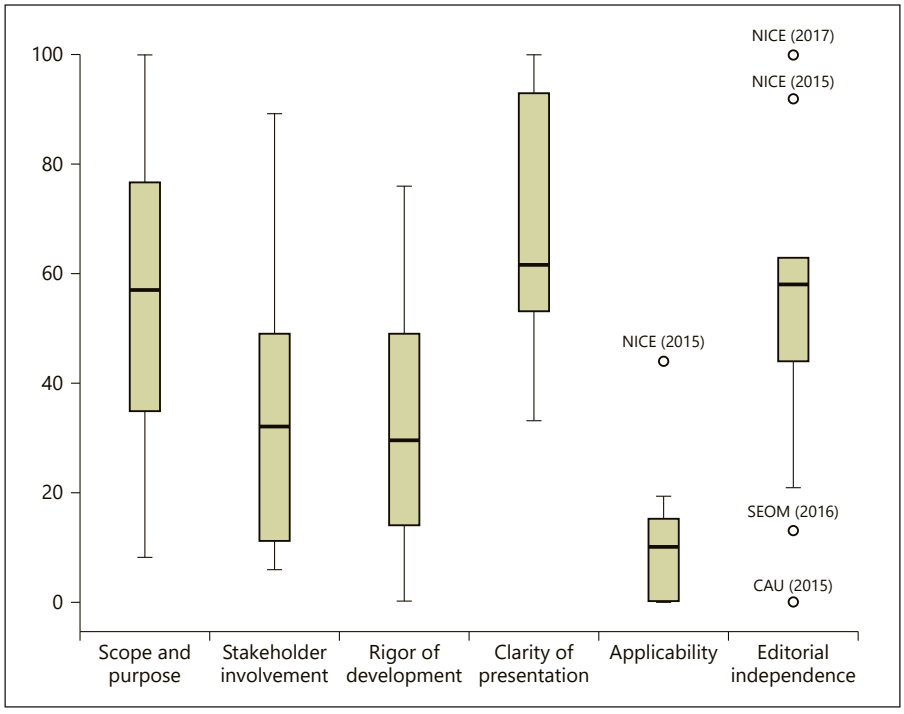
Results of Systematic Search

A total of 750 publications were identified by systematic literature search. One publication was found by screening the reference lists of the records found. After the removal of duplicates (n = 14), 736 publications were screened by analyzing the heading and abstract. Finally, 27 relevant publications remained for the review by full text and 16 publications met inclusion criteria for detailed analysis (Fig. 1).

Content of Included CPGs

The content and the topics of the analyzed CPGs differ significantly (Table 1). Four out of 16 CPGs include both the subjects of NMIBC and MIBC [23–26]. The topic NMIBC is exclusively addressed by 3 CPGs [16, 18, 19] and MIBC by 4 CPGs [15, 17, 20, 27]. Five guidelines deal with special aspects of bladder cancer, including robot-assisted radical cystectomy, laser technology, *Bacillus* Calmette-Guerin (BCG), and radiological diagnosis [12–14, 21, 22]. None of the guidelines include all aspects of bladder cancer (screening, pathology, diagnosis, treatment, and aftercare).

The identified CPGs were published by 11 organizations and institutions, respectively (Table 1). The EAU provides 4 guidelines with the topics NMIBC, MIBC, laser technologies for the treatment of NMIBC, and robotic-assisted radical cystectomy for the treatment of MIBC [19–22]. Pocket versions are available for NMIBC and



Color version available online

Fig. 2. Box plot of the AGREE II domains showing the median (black line within the box), the 25 and 75% quartile (box), the whiskers (1.5 interquartile range), and outliers (black circle). AGREE II, Appraisal of Guidelines for Research and Evaluation II.

MIBC. CPGs are updated annually; however, laser technologies and robotic-assisted surgery have both been discontinued in 2014. Except for the subdomain of screening, the EAU provides the most detailed CPGs for NMIBC and MIBC.

In the UK, a multidisciplinary team (MDT) composed of the British Uro-oncology group (BUG), the British Association of Urological Surgeons (BAUS) Section of Oncology, and the Action on Bladder Cancer (ABC) published their CPG on NMIBC and MIBC in 2013 [26]. The National Institute for Health and Care Excellence (NICE) guideline also covers both tumor entities with the latest update in 2015 [24]. Its evidence review provides an in-depth methodological insight spread out on nearly 1,000 pages with a quality assessment of the included studies using the GRADE (Grades of Recommendation Assessment, Development and Evaluation) instrument. A short version is also available. Further European guidelines on BC were published in 2014 by the European Society of Medical Oncology (ESMO) and in 2016 by the Spanish Sociedad Española de Oncología de Médica (SEOM) [23, 27].

In the USA, the NCCN, the American Urological Association (AUA), and the American College of Radiology (ACR) each published their own CPG on BC. The NCCN is a nonprofit alliance of 27 cancer centers providing a guideline on NMIBC and MIBC with the feature of graphical treatment algorithms [25, 30]. Their guideline is at least updated annually. The AUA established 2 CPGs for NMIBC and MIBC in cooperation with the Society of Urologic Oncology (SUO), the American Society of Clinical Oncology (ASCO), and the American Society for Radiation Oncology (ASTRO) [16, 17]. Focusing on imaging methods, the ACR addresses pretreatment staging and posttreatment surveillance of BC [12, 13].

The Canadian Urological Association (CUA) published a CPG exclusively for NMIBC focusing on the aspect of treatment [18]. The Alberta Health Services is providing a CPG for MIBC, yet its guideline for NMIBC is specifically for the treatment of BCG [14, 15].

Appraisal

The results of the CPG appraisal using the AGREE II instrument are shown in Table 3. All guidelines were independently evaluated by 2 appraisers. The estimated reliability (intraclass correlation coefficient) between both raters is 0.991 (95% CI: 0.973–0.997).

The overall evaluation was carried out in accordance with the predefined criteria. Five CPGs were "strongly recommended" due to AGREE II criteria (AUA NMIBC, EAU NMIBC, EAU MIBC, NICE, and NCCN), and 4 guidelines were "weakly recommended" (Alberta NMICB,

Alberta MIBC, AUA MIBC, and EAU Laser). Because of poor methodological quality, 7 CPGs were "not recommended" (ACR pretreatment, ACR posttreatment, CUA, EAU Robotic, ESMO, MDT, and SEOM).

Among the 6 domains of AGREE II in all guidelines, 3 median scores were rated ≥50% (Fig. 2). Overall clarity of presentation rated best with a median score of 61.5%. Thirteen CPGs revealed a domain score of ≥50%. For the editorial independence (median score 58.0% in all guidelines), 12 guidelines were rated ≥50%. In this domain, the EAU guidelines had a score of 63% compared with the NICE and NCCN guidelines with 92 and 100%, respectively. This was due to the fact that the EAU did not provide detailed information on the conflicts of interest for their guideline members during the draft of this manuscript. Meanwhile, precise information is given on the EAU's homepage. This improvement points out that the awareness of defining these facts is rising and would lead to a rating of 100% in this domain for the EAU guidelines, which is comparable with the NCCN guideline. At least this new aspect would not change the status of these "strongly recommended" guidelines. The median score of *scope and purpose* was 57.0% in all guidelines with 8 CPGs scoring ≥50%.

Three out of 6 domain median scores were rated ≤50%. For the domain of *stakeholder involvement* (median score 32.0% in all guidelines), 4 CPGs were rated ≥50%. Only 2 guidelines were scored ≥50% for *rigor of development* (median score 29.5% in all guidelines). *Applicability* is the domain with the lowest median score of 10.0%. None of the guidelines scored ≥50%. Fifteen CPGs scored <20%. Only 1 guideline (NICE) was rated with 44.0%.

Some of the AGREE II domain scores point out the reason why CPGs were "strongly" versus "weakly"/"not recommended." All "strongly recommended" guidelines were rated ≥50% in the domain *scope and purpose*. The NICE guideline was rated with a score of 100% due to its explicit completion of this domain. The guideline describes its overall objective, explains the covered health questions, and mentions the population to whom the guideline is meant to apply for. In the domain *rigor of development*, almost all "strongly recommended" CPGs scored >50%, but none of the "weakly"/"not recommended" were rated with a score of ≥50%. Most of the "weakly"/"not recommended" CPGs fail to show that the guideline has been externally reviewed by experts prior to its publication or do not clearly describe the methods of formulating recommendations. The EAU guidelines for NMIBC and MIBC scored 97 and 94% in the domain *clarity of presentation* while covering a wide range of topics. This excellent score results due to the fact that all recommendations are specific and unambiguous as well as easy to identify. Furthermore, the recommendations give different options for the management due to the patient's health condition.

Discussion

A considerable number of guidelines for bladder cancer are available in the literature. This reflects a general trend as the number of published CPGs has been steadily increasing since the 1980s [31]. With a growing interest in using guidelines, numerous associations or governments developed their own guidelines. However, there is a significant difference in the quality of these CPGs published by different institutions.

In this study, we assessed the methodological quality of 16 CPGs for bladder cancer using the AGREE II instrument. As a major result of our study, we were able to show the vast heterogeneity of the methodological quality between and within the published CPGs.

Differences in the Methodological Quality between Included Guidelines

First, there is a major discrepancy between the methodological quality of the CPGs developed by different institutions as only 5 (AUA NMIBC, EAU NMIBC, EAU MIBC, NICE, and NCCN) out of 16 CPGs were deemed to be "strongly recommended" [16, 19, 20, 24, 25]. The AUA NMIBC guideline includes the aspects of pathology, diagnosis, treatment, and aftercare. However, because of its extreme precision this CPG might only be of limited interest for clinicians in daily practice. Furthermore, the EAU NMIBC and MIBC guidelines cover a wide field of clinical aspects in a very detailed manner, aside from the topic of screening, which is not covered. Since the EAU guidelines are updated annually, they embody the predicate of a "living guideline." Being up to date and flexible will be one of the major challenges for relevant guidelines in the future, considering the rapid development of new therapies. The EAU guidelines have a well-arranged layout and are easy to use in daily clinical routine. A shortcoming of the EAU guidelines as well as other CPGs is the lack of a transparent assessment of the included literature. The NICE guideline is the most detailed and remarkable CPG due to the extraordinary financial and human resources. A very broad and detailed evidence review is available for this guideline to help decision-makers understand the quality of evidence, on

which the given recommendations were based on. The guideline authors used the GRADE (Grades of Recommendation Assessment, Development and Evaluation) instrument in the development of the guideline. This is considered to be a modern and valid instrument for assessing the methodological quality of included studies [32]. Therefore, the NICE guideline has the most solid methodological fundament of all reviewed guidelines. However, in contrast to the EAU guidelines, the update interval of the NICE guideline is much longer, which may reduce its usefulness in clinical routine. Another approach is used by the NCCN guideline: it is based on treatment algorithms presented as diagrams and is overall very comprehensive, which simplifies its usability. Deficiencies of this guideline are the missing external review process by experts prior to its publication and the intransparent assessment of the included literature. Finally, this CPG is 1 out of 5 "strongly recommended" guidelines.

Overall, 4 CPGs can only be "weakly recommended" [14, 15, 17, 21]. All of these CPGs showed a significant lack of methodological quality. Because of these limitations, the CPGs should be of limited interest in daily routine.

Differences between the AGREE II Domains within Included Guidelines

There is a heterogeneity in the score reached in different AGREE II domains within all analyzed CPGs. These domains refer to different aspects of guideline development (e.g., *scope and purpose*, *rigor of development*, and *clarity of presentation*, see Introduction). This heterogeneity illustrates that almost all guidelines disregard different aspects which are relevant for the methodological quality. Overall, the median scores of the domains vary between 10.0 and 61.5%. The domains *clarity of presentation* (61.5%), *editorial independence* (58.0%), and *scope and purpose* (57.5%) were rated best. In contrast, the domain *applicability* only scored a median of 10.0%. In particular, domains with a high methodological claim (*stakeholder involvement*, *rigor of development*, and *applicability*) had a lower score. These differences highlight the need of a qualified methodological support during the guideline development.

Transparency, independence, and disclosure are key issues in the developing process of trustable CPGs. The domain of *editorial independence* addresses a potential impact on CPG's content and recommendations by the funding body or conflicts of interest. Most publishing institutions of CPGs are aware of its importance. All "strongly recommended" CPGs disclose conflicts of interest of the guideline members in a more or less detailed fashion. Only 1 CPG (CUA) did not provide any information on this topic. None of the CPGs provide information about a selective exclusion of panel members on special topics in the event of a conflict of interest.

With a median score of 10.0%, the domain *applicability* received the lowest rating of all domains. This domain mainly addresses tools and barriers for implementation, monitoring/auditing criteria, as well as cost factors. Our assessment shows that none of the CPGs include sufficient information for clinical implementation. Moreover, most guidelines do not address methods to monitor implementation. Only 1 CPG (NICE) scored >20% due to its detailed cost analysis. Without knowing the barriers or feasibilities (e.g., costs for a recommended procedure) for the implementation of a CPG, the use of its recommendations may be hindered. This may reduce their meaning in daily clinical use.

The domain *rigor of development* may be seen as the best indicator for the methodological quality of a CPG as it covers all aspects of its systematic developing process such as systematic literature research, evidence selecting process, linking between evidence and recommendations, external review, and updating process. The median value of this domain was 29.5%, which can be considered weak, with a range between 7 and 76%. Only 2 CPGs (EAU NMIBC and NICE) were assessed with a score of ≥50%. Only a few CPGs clearly state their searching strategy, inclusion and exclusion criteria for evidence, or the interval for updating. In most CPGs, an external review is missing, which might decrease acceptance and quality. These results demonstrate the need for methodological experts or at least methodological experienced clinicians in the development of CPGs.

Differences within CPGs for NMIBC and MIBC

The AGREE II assessment has shown a methodological heterogeneity within the guideline topics for NMIBC and for MIBC. Three CPGs (AUA NMIBC, CUA, and EAU NIMBC) exclusively deal with the topic of NMIBC, one was "not recommended" and 2 were "strongly recommended." Four guidelines (Alberta MIBC, AUA MIBC, EAU MIBC, and SEOM) address the topic of MIBC only, one was "not recommended," 2 were "weakly recommended," and one "strongly recommended." Since all guideline organizations should have the same evidence base of literature, the heterogeneity reflects the different methodology of CPG development of individual organizations. It cannot be assessed to which extent

conflicts of interest affect heterogeneity since no guideline gives an exact indication of how these were considered in the preparation of recommendations.

National versus International CPGs

This systematic review identified CPGs with national (e.g., AUA and NICE) as well as international (e.g., EAU and ESMO) scopes. For national and international scopes, the mode of implementation and the quality criteria for implementation can be different. Furthermore, national and international guidelines differ in their source and amount of underlying funding. Another difference is that national guidelines are more suitable to consider and address costs of its own health care system. These are limitations of our study as CPGs with national and international scopes and funding can be of limited comparability. As an example, a cost discussion may be very meaningful for the NICE guideline in the homogenous health environment of the United Kingdom of Great Britain. On the other hand, a cost discussion may be much less meaningful for an international guideline such as the EAU guidelines, which cover a lot of diverse health care systems. This is also correct for the applicability. Different modes of implementation may result in different framework and adherence specifications for guideline users. Users may be tied to these specifications to different degrees. Because of this, recommendations of national guidelines may have much acceptance in guideline users and will result in a stronger adherence to the given recommendations."

Limitations of the AGREE II Instrument

As described above, the AGREE II instrument is a validated tool to assess the methodological quality of CPGs. Because of the assessment, guidelines are graded into recommended or not recommended guidelines. It has to be pointed out that the assessment does not make any statement of the contents' quality or the quality of the evidence the recommendations rely on. The content or the evidence of a methodologically poorly rated CPG does not have to be invalid automatically. In contrast, a recommended CPG may implicate an overrating of the evidence used. The guideline user must be aware of this bias and should not take an AGREE II assessment as an unchallenged quality title for the overall quality of a guideline. Furthermore, different health care systems and guideline-developing organizations will lead to different interpretations of data and influence medical recommendations, independent of the underlying methodological quality of guideline development.

Additional Value of the AGREE II Instrument

A methodology for developing CPGs, as proposed in the AGREE II tool, can be the basis of a reliable CPG the user can trust in. For this reason, the AGREE II instrument can be used as a checklist for the development of CPGs of high methodological quality. The process of a CPG compilation is time consuming and expensive, yet the publication and implementation may have massive influence on daily work and patient care as well as patients' lives and health systems. Because of this extensive procedure and its impact, the transparent development and documentation of the methodological compilation should become standard.

Conclusion

The presented quality appraisal of bladder cancer guidelines shows a vast heterogeneity of the methodological quality between the available guidelines. Considering the rapid development of new therapies, "living guidelines," such as the EAU NMIBC or MIBC guideline, will become more relevant in the future guideline's landscape. As CPGs of outstanding quality require tremendous financial and human resources, a cooperation between guideline organizations and governmental institutions might help to improve CPGs' methodological quality and limitations.

Conflict of Interest Statement

P.M., M.R., J.E.G., and S.C.S. were involved in the guideline development of the German S3-Guideline on bladder cancer, which was not analyzed in this publication.

Funding Sources

The authors did not receive any funding.

Author Contributions

P. Maisch: project development, systematic literature search, literature selection, quality assessment, data analysis, and manuscript writing. R. Retz: quality assessment and manuscript editing. J.E. Gschwend: project development and manuscript editing. F. Koll: literature selection and data analysis. S.C. Schmid: project development, literature selection, quality assessment, and manuscript writing.

References

1 Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;136(5):E359–86.
2 Global Burden of Disease Cancer Collaboration. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. JAMA Oncol. 2017;3(4):524–48.
3 Richters A, Aben KKH, Kiemeney LALM. The global burden of urinary bladder cancer: an update. World J Urol. 2019 Nov 1.
4 Institute of Medicine (US) Committee to Advise the Public Health Service on Clinical Practice Guidelines. Clinical practice guidelines: directions for a new program. Field MJ, Lohr KN, editors. Washington, DC: National Academies Press; 1990.
5 Burls A. AGREE II-improving the quality of clinical care. Lancet. 2010;376(9747):1128–9.
6 Merseburger AS, Apolo AB, Chowdhury S, Hahn NM, Galsky MD, Milowsky MI, et al. SIU-ICUD recommendations on bladder cancer: systemic therapy for metastatic bladder cancer. World J Urol. 2019;37(1):95–105.
7 Leow JJ, Bedke J, Chamie K, Collins JW, Daneshmand S, Grivas P, et al. SIU–ICUD consultation on bladder cancer: treatment of muscle-invasive bladder cancer. World J Urol. 2019;37(1):61–83.
8 Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting, and evaluation in health care. Prev Med. 2010;51(5):421–4.
9 Available from: http://www.agreetrust.org. [cited 2017 Dec 12].
10 AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care. 2003;12(1):18–23.
11 Available from: https://www.agreetrust.org/about-the-agree-enterprise/agree-history/. [cited 2017 Dec 12].
12 Leyendecker JR, Clingan MJ, Remer EM, Bishoff JT, Blaufox MD, Eberhardt SC, et al. ACR Appropriateness Criteria® pretreatment staging of invasive bladder cancer. Reston, VA: American College of Radiology; 2012.
13 Leyendecker JR, Clingan MJ, Eberhardt SC, Friedman B, Fulgham PF, Hartman MS, et al. ACR Appropriateness Criteria® post-treatment surveillance of bladder cancer. Reston, VA: American College of Radiology; 2014.
14 Alberta Health Services. Nonmuscle invasive bladder cancer. 2013.
15 Alberta Health Services. Muscle invasive and locally advanced/metastatic bladder cancer. 2013.
16 Chang SS, Boorjian SA, Chou R, Clark PE, Daneshmand S, Konety BR, et al. Diagnosis and treatment of non-muscle invasive bladder cancer: AUA/SUO Guideline. J Urol. 2016;196(4):1021–9.
17 Chang SS, Bochner BH, Chou R, Dreicer R, Kamat AM, Lerner SP, et al. Treatment of non-metastatic muscle-invasive bladder cancer: AUA/ASCO/ASTRO/SUO Guideline. J Urol. 2017;198(3):552–9.
18 Kassouf W, Traboulsi SL, Kulkarni GS, Breau RH, Zlotta A, Fairey A, et al. CUA guidelines on the management of non-muscle invasive bladder cancer. Can Urol Assoc J. 2015;9(9–10):E690–704.
19 European Association of Urology. Guidelines on non-muscle-invasive urothelial carcinoma of the bladder. 2018.
20 European Association of Urology. Guidelines on muscle-invasive and metastatic bladder cancer. 2018.
21 European Association of Urology. Guidelines on lasers and technologies. 2014.
22 European Association of Urology. Guidelines on robotic- and single-site surgery in urology. 2014.
23 Bellmunt J, Orsola A, Leow JJ, Wiegel T, De Santis M, Horwich A, et al. Bladder cancer: ESMO Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2014;25(Suppl 3):iii40–8.
24 National Institute for Health and Care Excellence. Bladder cancer: diagnosis and management of bladder cancer: © NICE (2015). BJU Int. 2015;120(6):755–65.
25 Flaig TW, Spiess PE, Agarwal N, Bangs R, Boorjian SA, Buyyounouski MK, et al. NCCN guidelines insights: bladder cancer, version 5.2018. J Natl Compr Canc Netw. 2018;16(9):1041–53.
26 British Uro-Oncology Group (BUG); British Association of Urological Surgeons (BAUS); Action on Bladder Cancer (ABC). Multi-disciplinary team (MDT) guidance for managing bladder cancer. 2013.
27 Lazaro M, Gallardo E, Doménech M, Pinto Á, González-del-Alba A, Del Alba AG, et al. SEOM clinical guideline for treatment of muscle-invasive and metastatic urothelial bladder cancer (2016). Clin Transl Oncol. 2016;18(12):1197–205.
28 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Open Med. 2009;3(3):e123-30.
29 Available from: https://www.agreetrust.org/resource-centre/ [cited 2017 Dec 12].
30 Bladder cancer: diagnosis and management of bladder cancer: © NICE (2015) Bladder cancer: diagnosis and management of bladder cancer. BJU Int. 2017;120(6):755–65.
31 Alonso-Coello P, Irfan A, Solà I, Gich I, Delgado-Noguera M, Rigau D, 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(6):e58.
32 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.