



A critical appraisal of urolithiasis clinical practice guidelines using the AGREE II instrument

Bangyu Zou^{#^}, Yuhao Zhou[#], Zhiqing He, Xiaofeng Zhou, Sicheng Dong, Xiaopeng Zheng, Ran Xu, Xiaolu Duan, Guohua Zeng

Department of Urology, Guangdong Key Laboratory of Urology, Minimally Invasive Surgery Center, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

Contributions: (I) Conception and design: B Zou, G Zeng; (II) Administrative support: X Duan, G Zeng; (III) Provision of study materials or patients: X Duan, G Zeng; (IV) Collection and assembly of data: B Zou, Y Zhou, Z He, X Zhou, S Dong, X Zheng, R Xu; (V) Data analysis and interpretation: B Zou, Yu Zhou, X Duan, G Zeng; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Guohua Zeng, MD, PhD. Department of Urology, Guangdong Key Laboratory of Urology, Minimally Invasive Surgery Center, The First Affiliated Hospital of Guangzhou Medical University, No. 1 Kangda Road, Haizhu District, Guangzhou 510230, China.

Email: gzgyzh@vip.sina.com.

Background: The Appraisal of Guidelines for Research and Evaluation II (AGREE II) instrument was developed to improve the methodological quality of clinical practice guidelines (CPGs). High-quality guidelines can provide reliable recommendations for different clinical issues. Currently, there is no quality appraisal of CPGs for urolithiasis. This study evaluated the quality of evidence-based CPGs for urolithiasis and provided new insights into improving guideline quality on urolithiasis.

Methods: Systematic reviews were conducted to identify urolithiasis CPGs in PubMed, electronic databases, and websites of medical associations from January 2009 to July 2022. The quality of included CPGs was evaluated by four reviewers using the AGREE II instrument. Subsequently, the scores of all domains in the AGREE II instrument were calculated.

Results: A total of 19 urolithiasis CPGs were identified for review: seven from Europe, six from USA, three from international union, two from Canada, and one from Asia. The agreement among reviewers was rated good [intraclass correlation coefficient (ICC), 0.806; 95% CI: 0.779–0.831]. The domains with the highest scores were scope and purpose (69.7%, 54.2–86.1%) and clarity of presentation (76.8%, 59.7–90.3%). The domains of stakeholder involvement (44.9%, 19.4–84.7%) and applicability (48.5%, 30.2–72.9%) gained the lowest score. Only five guidelines (26.3%) were considered “strongly recommended”.

Conclusions: The overall quality of the eligible CPGs was relatively high; however, future work is still needed in the domains of rigor of development, editorial independence, applicability, and stakeholder involvement.

Keywords: Urolithiasis; clinical practice guideline (CPG); Appraisal of Guidelines for Research and Evaluation II (AGREE II); quality assessment

Submitted Dec 18, 2022. Accepted for publication May 10, 2023. Published online Jun 13, 2023.

doi: 10.21037/tau-22-846

View this article at: <https://dx.doi.org/10.21037/tau-22-846>

[^] ORCID: 0000-0002-0453-1680.

Introduction

Urolithiasis is a common urological disease with prevalence rates ranging from 1% to 20% worldwide (1). Its economic burden is expected to increase by a further \$1.24 billion per year by 2030 (2). Calcium oxalate stones account for approximately 70% of kidney stones and are usually relevant to hypercalciuria or hyperoxaluria (3). There is also evidence that the incidence of certain stone types may be increased, such as uric acid stones (4). Accordingly, the recurrence rate of urinary stone is 30–50% within 10 years (5,6). Evidence is also emerging that there may be a link between nephrolithiasis and chronic kidney disease (7).

In accordance with Institute of Medicine, clinical practice guidelines (CPGs) are evidence-based documents that are systematically developed to help practitioners and patients in making decisions regarding appropriate health care or treatment options for specific clinical circumstances (8). The value of CPGs is influenced by methodology during the guideline formulation process (9). However, due to the differences in regional economy, medical level and target population, the quality of CPGs formulated by different organizations varies greatly (10). As CPGs aim at providing high-quality, safe alternative care options, more and more urological organizations recognize the importance of standardization in guideline development. Over the last few decades, numerous urolithiasis CPGs were published by

medical associations and governments. As far as we know, no reporting on quality assessment of CPGs pertaining to urolithiasis has been reported.

The AGREE II instrument is an international framework produced in 2009 to evaluate the development of practice guidelines and the quality of reporting (11). It covers six domains, has been translated into different languages, and has been advocated by many health organizations. This study used the AGREE II instrument to appraise the reporting quality of eligible CPGs and identify factors limiting the improvement of urolithiasis CPGs.

Methods

Registration

The systematic review has been registered in the International Prospective Register of Systematic Reviews (PROSPERO; registration number: CRD42021286736, https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=286736).

Search strategy

Systematic reviews were performed in the following sources from January 2009 to July 2022: (I) Medline (via PubMed); (II) four CPG databases: National Institute for Health and Care Excellence (NICE), Agency for Healthcare Research and Quality (AHRQ), Guidelines International Network (GIN), Scottish Intercollegiate Guidelines Network (SIGN); (III) six expert organizations of guideline developers: European Association of Urology (EAU), American Urological Association (AUA), Canadian Urological Association (CUA), Urological Association of Asia (UAA), International Consultation on Urologic Diseases (ICUD), Chinese Urological Association (CUA). The comprehensive search strategy in MEDLINE (via PubMed) combined MeSH terms and free text search (Table S1). Additionally, CPGs published in Chinese were identified by reviews of three Chinese electronic databases: China National Knowledge Infrastructure (CNKI), VIP Database, and WANFANG database. The search strategy in websites of guideline developers used the following keywords: “urolithiasis”, “nephrolithiasis”, “ureterolithiasis”, “calculi”, “stone”, “kidney”, “urinary”, “guideline”, “guidance”, “guide”, “recommendation”, “consensus”, and “suggestion”.

Highlight box

Key findings

- Most of urolithiasis clinical practice guidelines (CPGs) published in the past decade were classified as recommended with modifications, and improvements are required for the methodological quality.

What is known and what is new?

- The Appraisal of Guidelines for Research and Evaluation II (AGREE II) instrument was used to improve the reporting quality of CPGs. High-quality CPGs can provide doctors and patients with reliable recommendations for better outcomes. In this study, the reporting quality of evidence-based CPGs for urolithiasis was assessed, and new insights for improving guideline quality in urolithiasis are provided.

What is the implication, and what should change now?

- CPG development groups should refer to tools such as AGREE II to improve the quality of CPGs and pay more attention to reporting the domains of rigor of development, editorial independence, applicability, and stakeholder involvement.

Inclusion and exclusion criteria

Inclusion criteria were as follows: (I) CPGs providing evidence-supported recommendations based on assessment of evidence; (II) CPGs developed by a medical association or an expert group; (III) CPGs containing recommendations regarding the management of urolithiasis; (IV) English or Chinese publications; (V) CPGs published between January 2009 and July 2022. Exclusion criteria were as follows: (I) editorials, letters, systematic reviews, or other literature explaining CPGs; (II) old versions or duplicates of CPGs; (III) CPGs limited to the use of a specific technology; (IV) CPGs published in non-English or non-Chinese languages.

Selection and evaluation of guidelines

All articles from the above-mentioned sources were exported to Endnote X9. After eliminating the duplicates, titles and full texts were independently browsed by two reviewers (B. Zou and Y. Zhou). Any discrepancies regarding eligibility were resolved through discussion with two additional experts (X. Duan, G. Zeng).

The reporting and development quality of each eligible CPGs were evaluated using the AGREE II instrument. This instrument was developed and validated by international guideline developers and researchers (AGREE Next Steps Research Consortium) to evaluate six domains of guideline development, followed by two overall items (11).

In this study, each item was evaluated by using a seven-point scale ranging from 1 (totally disagree) to 7 (totally agree), based on AGREE Reporting Checklist (Table S2). Domain scores (from 0% to 100%) were measured using the formula: (obtained score – minimal possible score)/(maximal possible score – minimal possible score) × 100% (11). These scores were followed by two items: overall assessment (from 1 = worst quality to 7 = best quality) and whether or not they would recommend the CPG for clinical practice. The domains of rigor of development and applicability were assigned double weight. A guideline with overall score of more than 60% was considered “strongly recommended”, “recommended with modifications” for score between 30% and 60%, and “not recommended” for scores less than 30% (12).

Four well-trained reviewers (B. Zou, Y. Zhou, Z. He and X. Zhou) independently evaluated each item, assigning a score of 1–7 points, and then the domain scores were calculated.

Statistical analysis

Three reviewers (X. Zheng, S. Dong, R. Xu) extracted the general characteristics of CPGs: title, date of publication, origin, version, developer, disease concerned, funding, and target population. Agreement among the four appraisers was evaluated using ICC with a 95% CI. ICC was rated as poor reliability (<0.40), moderate (0.40–0.75), and good (>0.75) (13). Statistical analyses were processed using Office Excel 2019 and SPSS version 25.0 (IBM, Armonk, NY, USA).

Results

Literature search

Totally, 189 references were identified by systematic search of databases and websites. Of these, seven references were excluded because of duplicates, and 146 references were removed due to irrelevance or other reasons after the abstract was being browsed. The remaining 36 references were reviewed in full text, and seventeen references were excluded due to irrelevance or lack of systematic literature search. Ultimately, 19 eligible CPGs were identified for evaluation (Figure 1).

CPG characteristics

Table 1 lists the general characteristics of the 19 included guidelines, all of which were published from 2014 to 2022. Of these guidelines, ten (52.6%) focused on urinary stone, seven (36.8%) focused on renal stone, and two (10.5%) focused on ureteral stone. Five (26.3%) were developed by EAU, six (31.6%) in the USA, two (10.5%) in Canada, and the rest in Asia, United Kingdom, Italy, and international union. The funding sources of twelve guidelines (63.2%) were described in detail while seven (36.8%) were not revealed. The grading system used to assess the quality of evidence differed among the eligible CPGs: seven (36.8%) used University of Oxford Centre for Evidence-Based Medicine Levels of Evidence grading system (OCEBM), five (26.3%) used Grading of Recommendations Assessment, Development and Evaluation (GRADE) system, five (26.3%) used self-modified grading system, two (10.5%) used mixed systems. No CPGs in Chinese met the inclusion criteria.

Quality assessment

Table 2 and Table 3 present the domain scores of all included

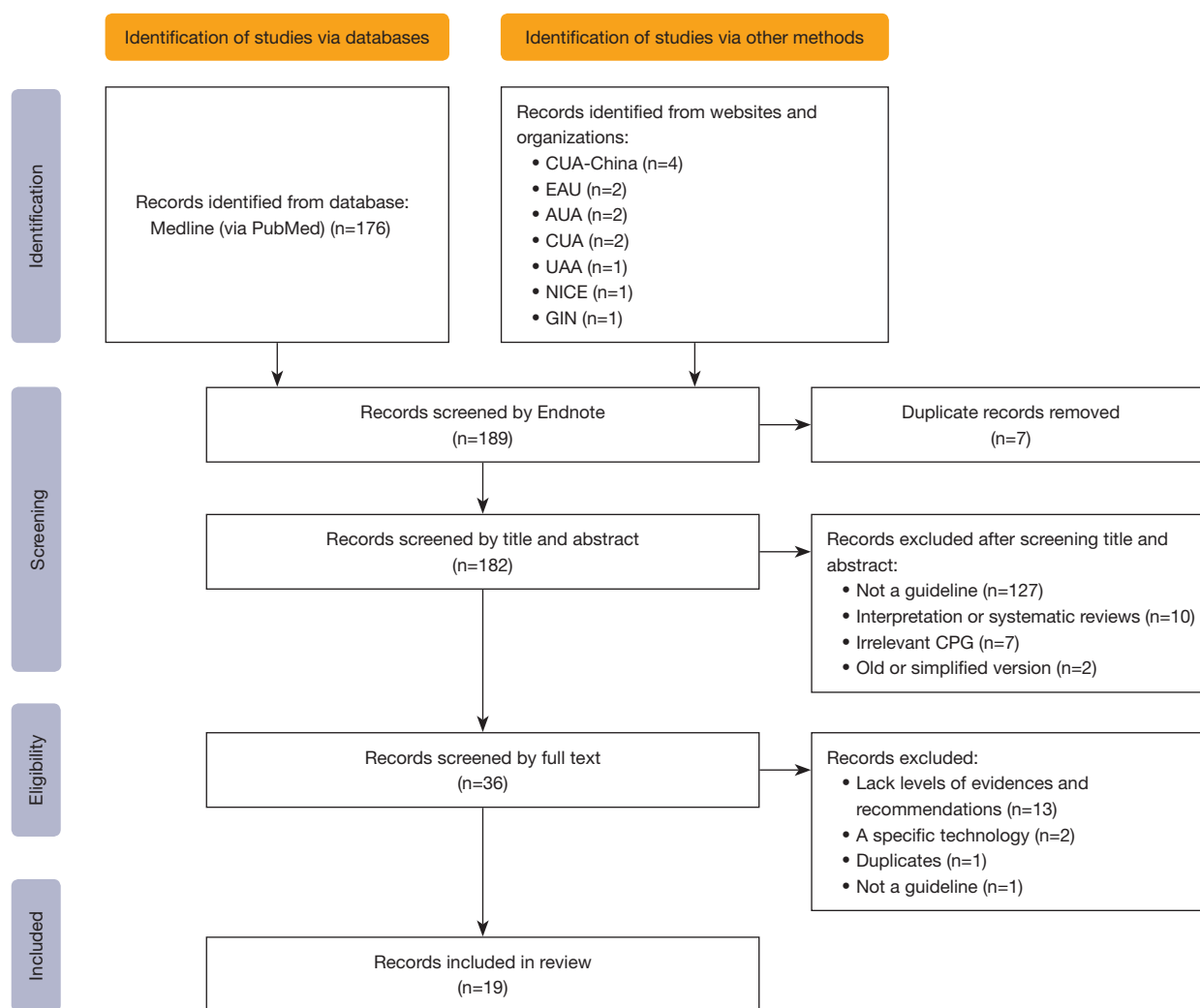


Figure 1 Flowchart of urolithiasis guideline search and selection. CUA, Canadian Urological Association; EAU, European Association of Urology; AUA, American Urological Association; UAA, Urological Association of Asia; NICE, National Institute for Health and Care Excellence; GIN, Guidelines International Network; CPG, clinical practice guideline.

CPGs and their overall recommendations. The ICC values among four appraisers was 0.806 (95% CI: 0.779–0.831), indicating high agreement. According to overall quality, five (26.3%) CPGs were rated as “strongly recommended” with overall scores >60%, while fourteen (73.7%) were considered as “recommended with modifications”, achieving scores of 30–60%. No guideline was considered as “not recommended” (Table 2).

Domain 1: scope and purpose

This domain addresses the main aims of the guidelines

and specific health issues. The scores of scope and purpose were relatively high (69.7%; 54.2–86.1%), with seventeen guidelines (89.5%) scoring above 60%, and all guidelines scoring more than 50%.

Domain 2: stakeholder involvement

This domain cares about the diversity of members in the guideline development team and preferences of the target population. The scores in this domain fluctuated dramatically (44.9%; 19.4–84.7%), with only three guidelines (15.8%) scoring more than 60%. Only two

Table 1 Characteristics of the identified guidelines

Title	Guideline	Origin	Development group	Subject	Type of urolithiasis	Version	Adult/children	Grading system	Funding
Dietary and pharmacologic management to prevent recurrent nephrolithiasis in adults: a clinical practice guideline from the American College of Physicians	Qaseem 2014 (14)	USA	ACP	Treatment	Renal stones	First	Adult	GRADE	ACP operating budget
Medical Management of Kidney Stones: AUA Guideline	Pearle 2014 (15)	USA	AUA	Diagnosis, prevention	Renal stones	First	Adult	Self-modified grading system	AUA
Surgical Management of Stones: American Urological Association/Endourological Society Guideline, PART I	Assimos 2016 (16)	USA	AUA, Endo	Diagnosis, prevention	Ureteral and renal stones	First	Both	Self-modified grading system	AUA and Endo
Surgical Management of Stones: American Urological Association/Endourological Society Guideline, PART II	Assimos 2016 (17)	USA	AUA, Endo	Treatment	Ureteral and renal stones	First	Both	Self-modified grading system	AUA and Endo
Dietary treatment of urinary risk factors for renal stone formation. A review of CLU Working Group	Prezioso 2015 (18)	Italy	CLU Working Group	Treatment	Renal stones	First	Both	GRADE	No report
Canadian Urological Association guideline: Management of ureteral calculi – Full-text	Lee 2021 (19)	Canada	CUA	Diagnosis, treatment	Ureteral stones	Updated	Both	OCEBM	No report
UPDATE – Canadian Urological Association guideline: Evaluation and medical management of kidney stones	Bhojani 2022 (20)	Canada	CUA	Diagnosis, treatment	Renal stones	Updated	Both	OCEBM	No report
Metabolic Evaluation and Recurrence Prevention for Urinary Stone Patients: EAU Guidelines	Skolarikos 2015 (21)	Europe	EAU	Diagnosis, treatment	Ureteral and renal stones	First	Both	OCEBM	None
EAU Guidelines on Interventional Treatment for Urolithiasis	Türk 2016 (22)	Europe	EAU	Treatment	Ureteral and renal stones	Updated	Both	OCEBM	None
EAU Guidelines on Diagnosis and Conservative Management of Urolithiasis	Türk 2016 (23)	Europe	EAU	Diagnosis, treatment	Ureteral and renal stones	Updated	Both	OCEBM	None

Table 1 (continued)

Table 1 (continued)

Title	Guideline	Origin	Development group	Subject	Type of urolithiasis	Version	Adult/children	Grading system	Funding
EAU Guidelines on Urolithiasis	Skolarikos 2022 (24)	Europe	EAU	Diagnosis, treatment	Ureteral and renal stones	Updated	Both	GRADE, OCEBM	EAU
EAU Guidelines on Paediatric Urology	Radmayr 2022 (25)	Europe	EAU	Diagnosis, treatment	Ureteral and renal stones	Updated	Children	GRADE, OCEBM	EAU
Urolithiasis: evaluation, dietary factors, and medical management: an update of the 2014 SIU-ICUD international consultation on stone disease	Jung 2017 (26)	International	SIU-ICUD	Diagnosis, treatment	Ureteral and renal stones	Updated	Both	OCEBM	No report
NICE Guideline – Renal and ureteric stones: assessment and management	NICE 2019 (27)	Britain	NICE	Diagnosis, treatment	Ureteral and renal stones	Updated	Both	GRADE	NICE
Metabolic diagnosis and medical prevention of calcium nephrolithiasis and its systemic manifestations: a consensus statement	Gambaro 2016 (28)	International	Expert committee	Diagnosis, treatment	Renal stones	First	Both	Self-modified grading system	No report
a-Blockers for uncomplicated ureteric stones: a clinical practice guideline	Vermandere 2018 (29)	International	Expert committee	Treatment	Ureteral stones	First	Both	GRADE	None
Evaluation and Medical Management of Patients with Cystine Nephrolithiasis: A Consensus Statement	Eisner 2020 (30)	USA	Expert committee	Diagnosis, treatment	Renal stones	First	Both	GRADE	None
Management of Nephrolithiasis in Pregnancy: Multi-Disciplinary Guidelines From an Academic Medical Center	Lee 2021 (31)	USA	Expert committee	Diagnosis, treatment	Renal stones	First	Adult	Self-modified grading system	No report
The Urological Association of Asia clinical guideline for urinary stone disease	Taguchi 2019 (32)	Asia	UAA	Diagnosis, treatment	Ureteral and renal stones	First	Both	OCEBM	No report

ACP, American College of Physicians; AUA, American Urological Association; Endo, Endourological Society; CLU, Clinical Lithiasis Unit; CUA, Canadian Urological Association; EAU, European Association of Urology; SIU-ICUD, The Société Internationale d'Urologie-International Consultation on Urologic Diseases; NICE, National Institute for Health and Care Excellence; UAA, Urological Association of Asia; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; OCEBM, Oxford Centre for Evidence-Based Medicine.

Table 2 Quality scores (%) of included clinical practice guidelines using AGREE II

Guideline	Score of the six AGREE II domains (%)						Overall quality	Overall assessment
	Scope and purpose	Stakeholder involvement	Rigor of development	Clarity and presentation	Applicability	Editorial independence		
Qaseem 2014 (14)	65.3	34.7	56.3	62.5	30.2	54.2	48.7	Recommended with modifications
Pearle 2014 (15)	72.2	56.9	67.2	75.0	40.6	60.4	60.0	Recommended with modifications
Assimos 2016 (16)	70.8	29.2	45.8	84.7	37.5	62.5	51.7	Recommended with modifications
Assimos 2016 (17)	70.8	29.2	43.2	83.3	42.7	58.3	51.7	Recommended with modifications
Prezioso 2015 (18)	56.9	27.8	47.9	59.7	36.5	16.7	41.2	Recommended with modifications
Lee 2021 (19)	70.8	19.4	38.0	76.4	41.7	37.5	45.4	Recommended with modifications
Bhojani 2022 (20)	68.1	23.6	54.2	79.2	40.6	27.1	48.4	Recommended with modifications
Skolarikos 2015 (21)	73.6	44.4	53.6	83.3	53.1	70.8	60.7	Strongly recommended
Türk 2016 (22)	70.8	43.1	54.7	79.2	46.9	77.1	59.2	Recommended with modifications
Türk 2016 (23)	54.2	37.5	51.0	76.4	40.6	79.2	53.8	Recommended with modifications
Skolarikos 2022 (24)	73.6	61.1	74.0	90.3	62.5	72.9	71.4	Strongly recommended
Radmayr 2022 (25)	70.8	52.8	68.8	84.7	55.2	75.0	66.4	Strongly recommended
Jung 2017 (26)	63.9	34.7	34.4	75.0	41.7	47.9	46.7	Recommended with modifications
NICE 2019 (27)	86.1	84.7	78.1	81.9	72.9	81.3	79.5	Strongly recommended
Gambaro 2016 (28)	76.4	51.4	47.9	63.9	49.0	54.2	54.9	Recommended with modifications
Vermandere 2018 (29)	70.8	81.9	62.5	65.3	71.9	81.3	71.0	Strongly recommended
Eisner 2020 (30)	72.2	47.2	45.3	80.6	43.8	81.3	57.4	Recommended with modifications
Lee 2021 (31)	68.1	45.8	50.5	72.2	62.5	50.0	57.8	Recommended with modifications
Taguchi 2019 (32)	69.4	48.6	56.3	86.1	52.1	43.8	58.1	Recommended with modifications

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

Table 3 AGREE II scores of the guidelines

Domain	Mean \pm SD (%)	Range (%)	Score <30%, n (%)	30% \leq Score \leq 60%, n (%)	Score >60%, n (%)
Scope and purpose	69.7 \pm 6.8	54.2–86.1	0 (0.0)	2 (10.5)	17 (89.5)
Stakeholder involvement	44.9 \pm 17.7	19.4–84.7	5 (26.3)	11 (57.9)	3 (15.8)
Rigor of development	54.2 \pm 11.7	34.4–78.1	0 (0.0)	14 (73.7)	5 (26.3)
Clarity and presentation	76.8 \pm 8.7	59.7–90.3	0 (0.0)	1 (5.3)	18 (94.7)
Applicability	48.5 \pm 11.9	30.2–72.9	0 (0.0)	15 (78.9)	4 (21.1)
Editorial independence	59.6 \pm 19.1	16.7–81.3	2 (10.5)	7 (36.8)	10 (52.6)

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

(10.5%) disclosed the involvement of methodologist.

Domain 3: rigor of development

This domain investigates the method used to review the evidence and the process to formulate recommendations. The domain of Rigor of development gained a mean score of 54.2% (34.4–78.1%), with five (26.3%) guidelines scoring above 60%. No guidelines scored below 30%.

Domain 4: clarity of presentation

This domain focuses on clarity of reporting and organization of recommendations. This domain received the highest score (76.8%; 59.7–90.3%) with eighteen (94.7%) guidelines scoring above 60%. The lowest score was 59.7% for dietary treatment guidelines for renal stone from CLU Working Group.

Domain 5: application

This domain is concerned with barriers, facilitators to its application, and advice or tools on how the recommendations can be put into practice. The mean score for Application was 48.5% (30.2–72.9%), among which four guidelines scored more than 60% as they reported tools and resources to facilitate implementation.

Domain 6: editorial independence

This domain addresses whether funding and conflicts of interest are properly documented. The mean score for Editorial independence was 59.6% (16.7–81.3%), with two guidelines (10.5%) given a score <30%. CLU Working Group guidelines for renal stone got the lowest

score of 16.7%.

Classification of CPGs quality

Table 4 presents the domain scores among different subgroups. Specifically, in the domain of clarity and presentation, the scores of CPGs that focused on renal stones and ureteral stones were higher than their counterparts that emphasized solely on renal stones or ureteral stones ($P<0.01$). Furthermore, in the domain of clarity, it was found that CPGs developed by medical specialty societies exhibited a slight increase in comparison to those developed by other organizations ($P<0.05$). Throughout the development of CPGs, an intriguing phenomenon was observed whereby there was a discernible elevation in AGREE II scores over time, particularly in the domain of Applicability ($P<0.05$). No significant differences were found with regards to the region of development group or the purpose for recommendation, nor were there any significant disparities observed between the first version and its subsequent updates.

Discussion

AGREE II assessment

This is the first research to adopt the AGREE II instrument to thoroughly assess the methodological quality of urolithiasis guidelines published in the past decade. However, this research was an analysis of the methodological quality of the included guidelines and did not compare and validate specific recommendations in the guidelines. Hence, comparison on the differences and consensus among CPGs in different regions are needed in future studies.

Table 4 AGREE II domain scores of guidelines according to different subgroups (mean \pm SD)

Subgroup	Scope and purpose	Stakeholder involvement	Rigor of development	Clarity and presentation	Applicability	Editorial independence	Overall
Region							
North America	69.79 \pm 2.43	35.76 \pm 13.00	50.07 \pm 9.11	76.74 \pm 7.11	42.45 \pm 9.14	53.91 \pm 16.42	52.65 \pm 5.22
Europe	69.44 \pm 10.85	50.20 \pm 18.55	61.16 \pm 12.15	79.37 \pm 9.70	52.53 \pm 12.61	67.56 \pm 22.72	61.74 \pm 12.39
International area+Asia	70.14 \pm 5.13	54.17 \pm 19.90	50.26 \pm 12.16	72.57 \pm 10.29	53.65 \pm 12.91	56.77 \pm 16.87	57.68 \pm 10.10
Type of urolithiasis discussed							
Renal stones/ureteral stones	68.98 \pm 5.51	43.21 \pm 19.49	52.20 \pm 8.94	70.52 \pm 7.79**	46.30 \pm 13.09	51.39 \pm 22.00	53.89 \pm 9.00
Renal stones+ureteral stones	70.42 \pm 7.99	46.53 \pm 16.85	55.99 \pm 13.91	82.50 \pm 4.59	50.52 \pm 11.04	66.88 \pm 13.19	59.92 \pm 10.04
Development group							
Medical specialty societies	68.80 \pm 5.25	39.64 \pm 12.86	53.65 \pm 11.66	79.70 \pm 7.03*	45.03 \pm 8.61	58.97 \pm 16.32	55.56 \pm 7.88
Others	71.76 \pm 9.61	56.48 \pm 22.34	55.38 \pm 12.68	70.60 \pm 9.19	56.08 \pm 15.24	60.76 \pm 25.93	60.32 \pm 13.35
Purpose for recommendation							
Diagnosis/prevention/treatment	68.25 \pm 5.46	43.25 \pm 19.97	53.94 \pm 8.86	72.82 \pm 10.26	43.75 \pm 13.46	58.63 \pm 21.02	54.79 \pm 9.57
Diagnosis/prevention+treatment	70.60 \pm 7.53	45.95 \pm 17.11	54.34 \pm 13.40	79.17 \pm 7.03	51.30 \pm 10.50	60.07 \pm 18.86	58.38 \pm 10.10
Year of publication							
2009–2016	67.90 \pm 7.62	39.35 \pm 10.37	51.97 \pm 7.14	74.23 \pm 9.75	41.90 \pm 6.98*	59.26 \pm 18.46	53.56 \pm 6.21
2017–2022	71.39 \pm 5.83	50.00 \pm 21.72	56.20 \pm 14.74	79.17 \pm 7.29	54.48 \pm 12.53	59.79 \pm 20.67	60.21 \pm 11.59
Version							
First	69.70 \pm 5.12	45.20 \pm 15.72	52.41 \pm 7.57	74.24 \pm 9.98	47.25 \pm 12.16	57.58 \pm 18.10	55.76 \pm 7.63
Updated	69.79 \pm 8.99	44.62 \pm 21.29	56.64 \pm 16.00	80.38 \pm 5.11	50.26 \pm 12.14	62.24 \pm 21.38	58.85 \pm 12.54

Data are presented as mean \pm SD. *, $P < 0.05$; **, $P < 0.01$. AGREE II, Appraisal of Guidelines for Research and Evaluation II.

Notably, although this study limited the language of guidelines to English and Chinese, no CPGs in Chinese met the inclusion criteria. All four guidelines written in Chinese that were included in the screening process were rejected because they failed to provide evidence-based recommendations. It has to be admitted that the standard framework for guideline development, such as AGREE II instrument and G-I-N Standards (33), has not been promoted and applied in the field of urolithiasis in China.

This analysis indicated that the domain of Clarity of presentation (76.8%) and the domain of Scope and purpose (69.7%) performed best, describing the objectives, health issues, and target populations. The recommendations of these guidelines were also mainly summarized in a box, in bold, underlined or as flow charts or algorithms. The presentation of recommendations would affect the usage of doctors and patients. A concise presentation would allow users to quickly locate the information they need,

while ambiguous guidelines would hinder the promotion and application of guidelines. This may be because guideline developers have focused on the professional aspects and neglected the simplicity and accessibility of recommendations.

The domain of Stakeholder involvement and the domain of Applicability obtained the lowest scores (<50%). Most guidelines failed to involve methodologists and patients in the development of guidelines. We have to emphasize again the irreplaceability of methodologists for this domain and even the entire guideline development. The participation of methodologists can maximize the integrity and standardization of CPGs, while the participation of patients is conducive to the emergence of patient-centered decision-making. As for the domain of Applicability, many CPGs neglected to consider barriers and facilitators to its application and missing the suggestions on how to apply the recommendations. Given the importance of this domain

in translating recommendations into clinical practice, guideline developers need to conduct pilot tests prior to publication.

Score on the domain of Editorial independence was not that high (59.6%), with many CPGs lacking an explicit statement about the external funding and its influence on the recommendations of the guideline. Independence of recommendation formation plays an important role in the reliability and transparency of the guideline. All strongly recommended CPGs reported conflicts of interest of the development panels in some extent. Finally, poor performance (54.2%) in the domain of rigor of development is of particular concern as it is regarded as the most important domain for methodological quality of guidelines and affect the development of the final recommendations (34).

Recommendations to improve guideline quality

In certain regions, advanced medical equipment and technology may not be available, leading to a lack of certain treatment options or diagnostic tools. On the other hand, differences in diet and lifestyle may result in a higher incidence of specific types of stones among urolithiasis patients in certain regions, and different subtypes of urolithiasis may require distinct treatment approaches. Therefore, to ensure the applicability and effectiveness of clinical practice guidelines in different regions, it is necessary to consider these regional differences and propose optimal treatment plans that correspond to different medical conditions, thereby facilitating the application of the guidelines.

Evidence-based recommendations are the cornerstone of contemporary guidelines, and we emphasize the need for transparent grading of evidence and recommendations to establish user trust. This research has revealed that numerous guidelines fail to disclose conflicts of interest, which can significantly impact the credibility of the guidelines. Thus, we propose that CPG development groups transparently disclose any potential conflicts of interest during the guideline development process. To ensure successful implementation of guidelines, we recommend prioritizing certain areas and soliciting user feedback to identify any barriers. Additionally, CPG development groups should be familiar with tools such as AGREE II, G-I-N Standards (33), Guidelines 2.0 (35) and incorporate them into the guidelines. Besides, guidelines should be reviewed for quality standards by methodologists before peer review. Journal editors should set higher

standards for publication, and only guidelines that satisfy the criteria for development can be published, which will in turn promote the spread of normative frameworks for guideline development. Finally, developers should pay more attention to controversial recommendations in different countries and regions which would lead to higher quality evidence and minimize duplication of effort.

Strengths and limitations of the study

This study had some limitations. Firstly, it focused on assessing the methodological quality of CPGs and did not compare the accuracy and feasibility of specific recommendations. Further research should be conducted to confirm consensus and differences between CPGs. Secondly, selection bias may exist because this study was limited to CPGs published in English or Chinese and excluded literature published in other forms (e.g., books, booklets, or government documents). Therefore, the results of this study may not necessarily apply to all CPGs for urolithiasis published globally. Finally, some supplemental materials may have been ignored in this study; consequently, in some cases, the quality of the guidelines may have been underestimated.

Nonetheless, the findings of this study remained reliable. Our team included clinical specialist with experience in urolithiasis CPGs development. Systematic literature searches were conducted, and significant agreement among reviewers was obtained, which ensured the reliability of our conclusions. Additionally, the AGREE II instrument was used to evaluate the reporting quality of eligible CPGs, which is validated by several international health organizations and applied to the development of CPGs worldwide.

Conclusions

This study aimed to evaluate urolithiasis CPGs in an effort to improve the reporting quality of CPGs and patients outcomes for those suffering from urolithiasis. Based on the AGREE II instrument, most of the included urolithiasis CPGs have a well-defined scope and purpose, as well as adequate clarity of presentation. However, future work is still needed in the domains of rigor of development, editorial independence, applicability, and stakeholder involvement.

Acknowledgments

Funding: This work was financed by grants from the

National Natural Science Foundation of China (Nos. 82070721 and 82270798), National Key Research and Development Program (No. 2020YFC2002700), Science and Technology Plan Project of Guangzhou (No. 202102010150 and No. 202002030316).

Footnote

Peer Review File: Available at <https://tau.amegroups.com/article/view/10.21037/tau-22-846/prf>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://tau.amegroups.com/article/view/10.21037/tau-22-846/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.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Raheem OA, Khandwala YS, Sur RL, et al. Burden of Urolithiasis: Trends in Prevalence, Treatments, and Costs. *Eur Urol Focus* 2017;3:18-26.
2. Antonelli JA, Maalouf NM, Pearle MS, et al. Use of the National Health and Nutrition Examination Survey to calculate the impact of obesity and diabetes on cost and prevalence of urolithiasis in 2030. *Eur Urol* 2014;66:724-9.
3. Xu H, Zisman AL, Coe FL, et al. Kidney stones: an update on current pharmacological management and future directions. *Expert Opin Pharmacother* 2013;14:435-47.
4. Chou YH, Su CM, Li CC, et al. Difference in urinary stone components between obese and non-obese patients. *Urol Res* 2011;39:283-7.
5. Uribarri J, Oh MS, Carroll HJ. The first kidney stone. *Ann Intern Med* 1989;111:1006-9.
6. Rule AD, Lieske JC, Li X, et al. The ROKS nomogram for predicting a second symptomatic stone episode. *J Am Soc Nephrol* 2014;25:2878-86.
7. Zhe M, Hang Z. Nephrolithiasis as a risk factor of chronic kidney disease: a meta-analysis of cohort studies with 4,770,691 participants. *Urolithiasis* 2017;45:441-8.
8. Field MJ, Lohr KN. *Clinical Practice Guidelines: Directions for a New Program*. Washington: Washington (DC): National Academies Press; 1990.
9. Grol R, Cluzeau FA, Burgers JS. *Clinical practice guidelines: towards better quality guidelines and increased international collaboration*. *Br J Cancer* 2003;89 Suppl 1:S4-8.
10. Jiang M, Liao LY, Liu XQ, et al. Quality Assessment of Clinical Practice Guidelines for Respiratory Diseases in China: A Systematic Appraisal. *Chest* 2015;148:759-66.
11. 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.
12. Jiang M, Guan WJ, Fang ZF, et al. A Critical Review of the Quality of Cough Clinical Practice Guidelines. *Chest* 2016;150:777-88.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
14. Qaseem A, Dallas P, Forcica MA, et al. Dietary and pharmacologic management to prevent recurrent nephrolithiasis in adults: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 2014;161:659-67.
15. Pearle MS, Goldfarb DS, Assimos DG, et al. Medical management of kidney stones: AUA guideline. *J Urol* 2014;192:316-24.
16. Assimos D, Krambeck A, Miller NL, et al. Surgical Management of Stones: American Urological Association/Endourological Society Guideline, PART I. *J Urol* 2016;196:1153-60.
17. Assimos D, Krambeck A, Miller NL, et al. Surgical Management of Stones: American Urological Association/Endourological Society Guideline, PART II. *J Urol* 2016;196:1161-9.
18. Prezioso D, Strazzullo P, Lotti T, et al. Dietary treatment of urinary risk factors for renal stone formation. A review of CLU Working Group. *Arch Ital Urol Androl* 2015;87:105-20.
19. Lee JY, Andonian S, Bhojani N, et al. Canadian Urological Association guideline: Management of ureteral calculi - Full-text. *Can Urol Assoc J* 2021;15:E676-90.
20. Bhojani N, Bjazevic J, Wallace B, et al. UPDATE - Canadian Urological Association guideline: Evaluation and

- medical management of kidney stones. *Can Urol Assoc J* 2022;16:175-88.
21. Skolarikos A, Straub M, Knoll T, et al. Metabolic evaluation and recurrence prevention for urinary stone patients: EAU guidelines. *Eur Urol* 2015;67:750-63.
 22. Türk C, Petřík A, Sarica K, et al. EAU Guidelines on Interventional Treatment for Urolithiasis. *Eur Urol* 2016;69:475-82.
 23. Türk C, Petřík A, Sarica K, et al. EAU Guidelines on Diagnosis and Conservative Management of Urolithiasis. *Eur Urol* 2016;69:468-74.
 24. EAU Guidelines on Urolithiasis [database on the Internet]2022. Available online: <https://uroweb.org/guidelines/urolithiasis>. Accessed: 26 July 2022.
 25. EAU Guidelines on Paediatric Urology [database on the Internet]2022. Available online: <https://uroweb.org/guidelines/paediatric-urology>. Accessed: 26 July 2022.
 26. Jung H, Andonian S, Assimos D, et al. Urolithiasis: evaluation, dietary factors, and medical management: an update of the 2014 SIU-ICUD international consultation on stone disease. *World J Urol* 2017;35:1331-40.
 27. NICE Guideline - Renal and ureteric stones: assessment and management: NICE (2019) Renal and ureteric stones: assessment and management. *BJU Int* 2019;123:220-32.
 28. Gambaro G, Croppi E, Coe F, et al. Metabolic diagnosis and medical prevention of calcium nephrolithiasis and its systemic manifestations: a consensus statement. *J Nephrol* 2016;29:715-34.
 29. Vermandere M, Kuijpers T, Burgers JS, et al. alpha-Blockers for uncomplicated ureteric stones: a clinical practice guideline. *BJU Int* 2018;122:924-31.
 30. Eisner BH, Goldfarb DS, Baum MA, et al. Evaluation and Medical Management of Patients with Cystine Nephrolithiasis: A Consensus Statement. *J Endourol* 2020;34:1103-10.
 31. Lee MS, Fenstermaker MA, Naoum EE, et al. Management of Nephrolithiasis in Pregnancy: Multi-Disciplinary Guidelines From an Academic Medical Center. *Front Surg* 2021;8:796876.
 32. Taguchi K, Cho SY, Ng AC, et al. The Urological Association of Asia clinical guideline for urinary stone disease. *Int J Urol* 2019;26:688-709.
 33. Qaseem A, Forland F, Macbeth F, et al. Guidelines International Network: toward international standards for clinical practice guidelines. *Ann Intern Med* 2012;156:525-31.
 34. Hoffmann-Eßer W, Siering U, Neugebauer EAM, et al. Guideline appraisal with AGREE II: online survey of the potential influence of AGREE II items on overall assessment of guideline quality and recommendation for use. *BMC Health Serv Res* 2018;18:143.
 35. Schünemann HJ, Wiercioch W, Etzeandia I, et al. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. *CMAJ* 2014;186:E123-42.

Cite this article as: Zou B, Zhou Y, He Z, Zhou X, Dong S, Zheng X, Xu R, Duan X, Zeng G. A critical appraisal of urolithiasis clinical practice guidelines using the AGREE II instrument. *Transl Androl Urol* 2023;12(6):977-988. doi: 10.21037/tau-22-846