

## ORIGINAL ARTICLE

# Do the clinical practice guidelines for paediatric dentistry meet the quality standards? A meta-research and quality appraisal using the AGREE II tool

Rokaia Ahmed Elagami<sup>1</sup>  | Caroline Mariano Laux<sup>1</sup> | Claudia López Gallegos<sup>1</sup> | Tamara Kerber Tedesco<sup>1</sup>  | Thais Gimenez Cóvos<sup>2</sup> | Mariana Minatel Braga<sup>1</sup> | Fausto Medeiros Mendes<sup>1</sup>  | Maximiliano Sérgio Cenci<sup>3</sup> | Daniela Prócida Raggio<sup>1,4</sup>

<sup>1</sup>Department of Paediatric Dentistry, School of Dentistry, University of São Paulo, São Paulo, Brazil

<sup>2</sup>School of Dentistry, Metropolitan University of Santos, São Paulo, Brazil

<sup>3</sup>Department of Dentistry, Radboud University Medical Center, Research Institute for Medical Innovation, Nijmegen, The Netherlands

<sup>4</sup>School of Dentistry, Cardiff University, Cardiff, UK

## Correspondence

Maximiliano Sérgio Cenci, Department of Dentistry, Radboud University Medical Center, Research Institute for Medical Innovation, Postbus 9101, 6500 HB Nijmegen, The Netherlands.  
Email: [Max.Cenci@radboudumc.nl](mailto:Max.Cenci@radboudumc.nl)

## Funding information

Conselho Nacional de Desenvolvimento Científico e Tecnológico; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Grant/Award Number: 001

## Abstract

**Background:** Clinical practice guidelines (CPGs) enhance health care and aid clinicians' decisions.

**Aim:** To evaluate the quality of clinical guidelines in paediatric dentistry using the AGREE II tool.

**Design:** PubMed, EMBASE, Scopus, LIVIVO, Lilacs, international guidelines websites, scientific societies, and gray literature were searched until September 2021. We included paediatric dental clinical guidelines and excluded drafts or guidelines for patients with special needs. Two independent reviewers performed quality assessment using the APPRAISAL OF GUIDELINES FOR RESEARCH & EVALUATION II (AGREE II) instrument. We calculated the mean overall domain scores (95% confidence interval) for each guideline. We used regression analysis to correlate the score of overall assessment and the six domains of AGREE II with guideline characteristics.

**Results:** Forty-four guidelines were included in this study. Highest mean score was for Domain 4 (Clarity of Presentation; 58%, 95% CI: 50.8–64.9), whereas the lowest was for Domain 5 (Applicability; 16%, 95% CI: 10.8–21.4). The reporting quality was improved in Domains 1–5 with reporting checklists ( $p < .001$ ), whereas that of Domain 6 was improved by decreasing years since publication ( $p = .047$ ).

**Conclusion:** Paediatric dental guidelines do not comply with the methodological quality standard, especially in Domain 5 (Applicability). The AGREE reporting checklist should be implemented with a system to evaluate the certainty of evidence for future guidelines.

## KEYWORDS

AGREE II tool, clinical practice guidelines, meta-research, paediatric dentistry, quality appraisal

## 1 | INTRODUCTION

Paediatric dentistry is an integrated field concerned with preventing, diagnosing, intercepting, and treating oral health problems, such as dental caries, periodontal problems, and tooth loss in children and adolescents.<sup>1</sup> These problems affect the growth and development process, by avoiding the socioeconomical and environmental risk factors (i.e., poor oral hygiene, malnutrition, lower source of fluoride, or poor diet) and implementing early-management programs to prevent and minimize oral health drawbacks, thus allowing the integral development from the maternal period through adolescence.<sup>2</sup> Henceforth, clinicians' decisions should rely on evidence-based best practices in the best interest of the patient's health.<sup>3</sup>

The American Dental Association (ADA) defined evidence-based dentistry (EBD) as an "approach to oral health care that requires the judicious integration of systematic assessments of clinically relevant scientific evidence, relating to the patient's oral and medical condition and history, with the dentist's clinical expertise and the patient's treatment needs and preferences."<sup>4</sup> Therefore, best evidence should originate from the highest quality and level of evidence (e.g., types of studies such as clinical trials, and systematic reviews).<sup>5</sup> To comply with the EBD, clinical practice guidelines (CPGs) have been developed, which are systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances. Hence, high-quality CPGs are considered as one of the gold standard sources for translating evidence to clinical practice, achieving cost-effective integration of patient care with clear recommendations.<sup>6</sup> Therefore, CPGs should be developed through rigorous methodologies that establish their quality. There are, however, potential biases or conflicts of interest when conducting a CPG that could lead to inappropriate recommendations based on insufficient or poor evidence, leading to potential patient harm.<sup>7</sup>

Several paediatric dental societies and organizations have published CPGs focusing on multiple topics in paediatric dentistry; some of them, however, might not be properly designed, and some recommendations are not evidence-based.<sup>8</sup> Moreover, not all the CPG developers use systems to evaluate the level of evidence, such as the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach, a grading system designed to rate the certainty of a body of evidence in systematic reviews and CPGs.<sup>9</sup>

Due to the increase in the numbers of published CPGs and recommendations, an international group from 13 countries decided to develop policies to evaluate the

### Why this paper is important to paediatric dentists

- The poor quality of the paediatric dentistry guideline reports may introduce biases and neglect patient preferences.
- Our results can help guide the development of more robust and reliable guidelines in the field.
- This paper serves as a call for action for paediatric dentists to critically evaluate the quality of guideline reports, advocate improved development processes, and utilize high-quality guidelines to enhance patient care in the field of paediatric dentistry.

quality of CPGs and ensure their high quality before endorsing them.<sup>10</sup> The APPRAISAL OF GUIDELINES FOR RESEARCH & EVALUATION II (AGREE II) instrument is a generic tool developed to ensure the quality of publishing guidelines and recommendations.<sup>10-12</sup> The use of a reporting checklist for CPGs recommended by the Enhancing the QUALity and Transparency Of health Research (EQUATOR) network is mandatory, whereas conducting CPGs ensures transparency, facilitates evidence presentation, and links recommendations to the strength of evidence.<sup>13,14</sup> The quality of CPGs in paediatric dentistry has been evaluated in a previous study that focused on CPGs published in English up to November 2007.<sup>8</sup> The study, however, used an outdated version of the AGREE tool to assess the reporting quality of these guidelines. Therefore, our goal was to provide an updated evaluation of the quality of published CPGs and recommendations in paediatric dentistry using the AGREE II tool.

## 2 | MATERIALS AND METHODS

Our study is a meta-research that aims to answer the research question: "Do the clinical practice guidelines on paediatric dentistry meet the quality standards as measured by the AGREE II instrument?" To ensure transparency and rigor, we registered our study protocol prospectively on the Open Science Framework (DOI <https://doi.org/10.17605/OSF.IO/BFNGW>).

### 2.1 | Deviation from the protocol

We planned in our registered protocol to include best practice documents, i.e., the documents that contain

best clinical or administrative practice, the evidence about what works for a situation/need/desire, and the resources available,<sup>15</sup> and intended to use the AGREE REX as a separate tool to evaluate them. The AGREE REX tool is a complementary tool to the AGREE II tool and can be used separately to evaluate the clinical credibility of the included recommendations.<sup>16</sup> Four evaluators (RAE, DPR, CML, and CLG), however, performed an appraisal assessment of two best practice documents, and we found that these documents lacked a transparent and declared methodology compared with the stated CPGs. After the discussion, we decided not to include the best practices in our study as they could lead to inaccurate results and deviate from our main aim of evaluating CPGs using the AGREE II tool. Moreover, we made the decision not to use the AGREE REX tool because the AGREE II tool already covers aspects related to recommendations, albeit in a broader context. By solely utilizing the AGREE II tool, we can avoid redundancy in the evaluation process, as it encompasses the assessment of recommendations along with other essential methodological aspects. We, however, recommend conducting further research to assess the reporting quality of the best practices document by employing a validated reporting quality assessment tool. This would provide additional insights into the quality of reporting within these documents. In addition, we did not calculate the intraclass correlation coefficient (ICC) for the scores between the reviewers. We considered the consistency instead of agreement and allowed the variation between the reviewers by up to 2 points, but if it varied by  $\geq 3$  points, a consensus was reached after a discussion with a third expert reviewer.

## 2.2 | Selection of the guidelines and recommendation

A literature search until September 2021 was conducted to identify the guidelines related to paediatric dentistry in the following databases: Medline (PubMed), Excerpta Medica Database (EMBASE), Scopus, LIVIVO, and Caribbean Health Sciences Literature (LILACS). In addition, Scientific societies and International Associations' websites were searched, such as the following: the British Society of Paediatric Dentistry—BSPD, American Academy of Pediatric Dentistry—AAPD, European Academy of Paediatric Dentistry—EAPD, Guidelines International Network—GIN, National Institute for Health and Care Excellence—NICE, Scottish Dental Clinical Effectiveness Programme—SDCEP, and ADA. Furthermore, the gray literature was searched through ProQuest and Turning Research Into

Practice database—TRIP—to identify the possibly eligible literature not identified during database searches. There were no restrictions on the language or the year of publication. The most recent updated version of CPG was considered.

The key terms used in the search strategy for this study were as follows: “Paediatric Dentistry,” “primary teeth and Permanent teeth,” “Guideline,” “clinical practice guideline,” “Recommendations” with terms from a controlled vocabulary (MeSH terms), keywords, synonyms, related terms, combined with Boolean operators “OR” and “AND” (Appendix S1).

## 2.3 | Eligibility criteria

### 2.3.1 | Inclusion criteria

- A. Clinical practice guidelines that contain a “statement” or “guideline” or provide “recommendation” directed to paediatric dentists; and
- B. Guidelines that included at least one recommendation related to paediatric dentistry based on the literature and expert opinion.

### 2.3.2 | Exclusion criteria

- A. Guidelines designed specifically for patients with special needs in paediatric dentistry;
- B. Protocols for the development of CPGs and draft of the CPGs; and
- C. Outdated guidelines when the most updated is retrieved.

Two independent reviewers (RAE and CLG) screened, evaluated the eligibility, and included the studies. Discrepancies were resolved through joint discussion with a third expert reviewer (DPR). The identified references were uploaded to the EndNote reference management program web (<https://www.myendnoteweb.com>) to remove the duplicates, and the remaining results were exported to Rayyan.<sup>17</sup> Rayyan, a free web tool (Beta), was used for storing the data for a long time, and thus facilitating the reviewers working on Phases 1 and 2. The references that were collected in Excel form were uploaded manually. For both methods, after duplicates were removed, documents were screened through Phase 1, and documents were analyzed based on the title and abstract. Documents that did not meet the inclusion criteria described above were eliminated. In Phase 2, a full-text read was performed to check the exclusion criteria.

## 2.4 | Data extraction and information handling

All the supplementary documents related to the CPGs were collected and analyzed. Two reviewers (RAE, CML) were responsible for extracting the following data from the included CPGs, independently: name of organization, the language, number of author(s) and organizations, whether it is single or multicenter, country, publication year, title, name of the journal, journal's impact factor according to Journal Citation Report (JCR) 2020, the system used for level of evidence and grading of recommendations, study's purpose, target population, development group, target users, systematic methods of evidence search, evidence strengths and limitations, recommendations methods, evidence eligibility criteria, benefits/side effects/risks, expert peer-reviewed, guideline update procedure, recommendation–evidence linkage, concrete/unequivocal recommendations, problem/disease management options, identifiable recommendations, application facilitators/barriers, how to put recommendations into advice, application of costs, monitoring/auditing criteria, content influence by funding, and development of group competing interests.

## 2.5 | Quality of evidence assessment and appraisal with AGREE II tool

Two reviewers (RAE, CML) completed the AGREE developer's recommended online training tool from [www.agreetrust.org](http://www.agreetrust.org). Then, we conducted a pilot evaluation on one CPG excluded from our analysis. Afterward, the two appraisers independently evaluated each guideline through the "MY AGREE PLUS" function according to the domain-based instrument AGREE instrument user manual.<sup>11,12</sup> The updated AGREE II instrument consists of 23 items grouped under six domains (Appendix S2). Each domain aims to assess each quality aspect of the guidelines, comprising: Scope and Purpose, Stakeholder Involvement, Rigour of Development, Clarity of Presentation, Applicability, and Editorial Independence. Furthermore, an overall assessment includes the rating of the overall quality of the guideline on a scale from 1 to 7, and whether we recommend the guideline for use, recommend with modification, or do not recommend it for use in practice.

Each item was rated on a Likert scale from 1 to 7, with 1 referring that the concept is very poorly reported (strongly disagree) and 7 referring that the full criteria are reported (strongly agree). The overall domain score was calculated by summing up all the scores of the items in that domain by the two reviewers (obtained score). The minimum

possible score was calculated for each domain by applying Formula 1 (strongly disagree) × (number of items for that domain) × (number of appraisers), whereas the maximum possible score was calculated for each domain by applying Formula 7 (strongly agree) × (number of items for that domain) × (number of appraisers).

The following equation was used to calculate the overall score for each domain<sup>12</sup>:

$$\frac{\text{Obtained score} - \text{Minimum possible score}}{\text{Maximum possible score} - \text{Minimum possible score}} \times 100.$$

As the overall assessment is individual in nature, the AGREE II manual did not declare a method to scrutinize the second question. Hoffmann-Eßer et al.<sup>18</sup> found that Domains 3, 4, and 5 had the most substantial influence on overall guideline quality. Therefore, we recommended a guideline for use when two of the three domains are with a score ≥60%, recommended for customized use with modification when two domains were between >30% and <60%, and did not recommend for use when two domains were ≤30%. If we found a guideline to receive equal distribution between the three domains, we decided according to the score of Domain 3 as it had the strongest influence.

## 2.6 | Statistical analysis

We assessed inter-rater reliability and agreement for eligibility criteria using Cohen's kappa test. Descriptive statistics were used to report the characteristics of the CPGs, as well as each assessor's total score and domain score. To calculate the domain score, we aggregated all item scores within the domain and transformed them into a maximum score of 100% for each item. We then calculated the mean overall score and 95% confidence interval (CI) for each domain. We used regression analysis (logistics and linear) to correlate the score of overall assessment and each domain of the six domains of the AGREE II tool (as the six domains cannot be aggregated into single score) with the following characteristics: years since publication, number of authors, number of centers, country, used reporting checklist, and used system to evaluate level of evidence. We conducted all statistical analyses using Stata/SE version 15.0 (StataCorp).

## 3 | RESULTS

### 3.1 | Guideline selection

The initial search identified 4906 documents, 129 full-text articles were screened for eligibility, and 44 guidelines were included in our analysis (Figure 1). The inter-rater

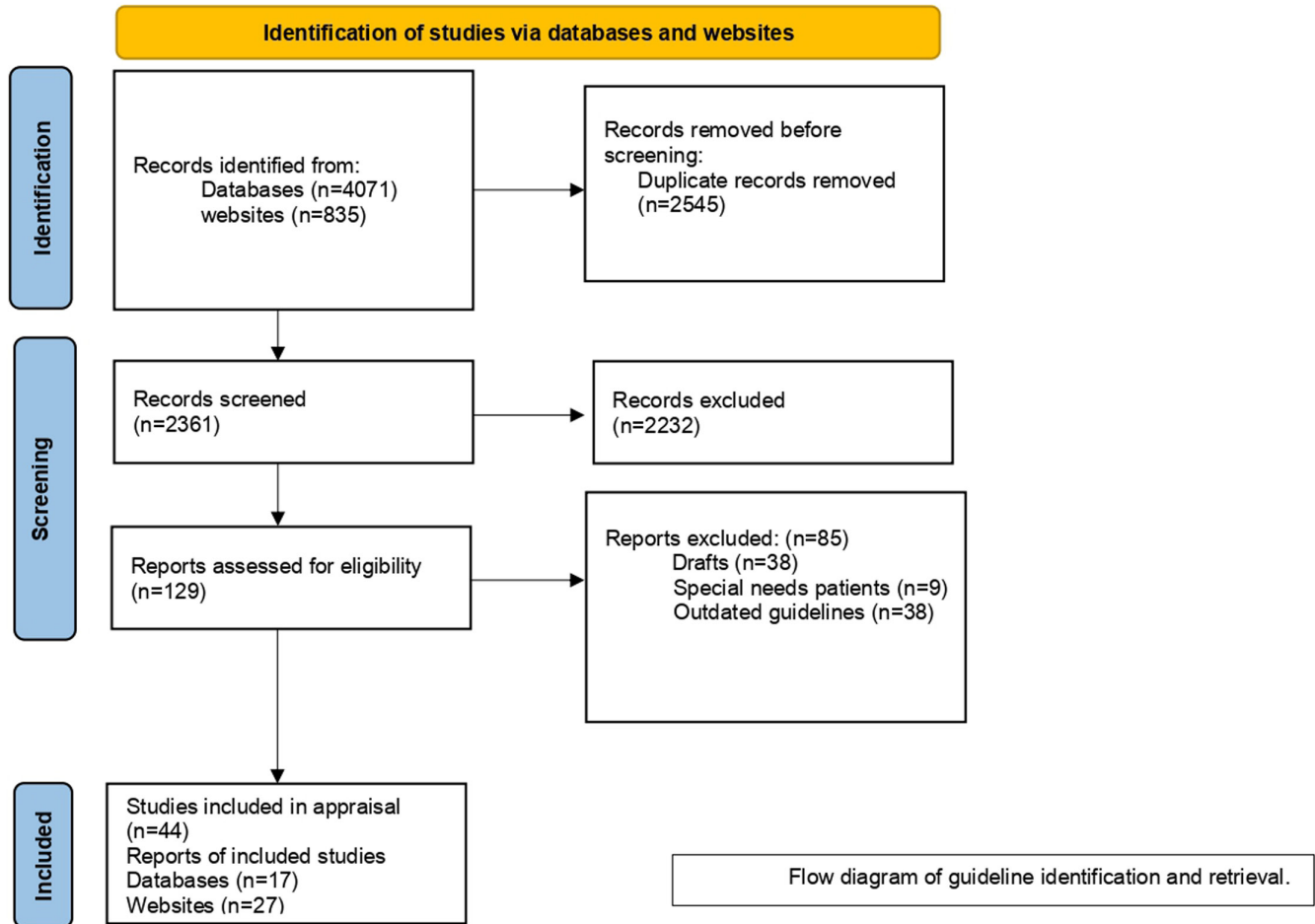


FIGURE 1 Flow diagram of guideline identification and retrieval.

reliability and agreement using the Cohen kappa coefficient was  $k=0.67$  (substantial agreement). Appendix S3 presents the title of the included guidelines with the organization's name.

### 3.2 | Guidelines demographics

Table 1 provides a general characteristic of the included CPGs. Eighteen were developed in North America (14 in the United States and four in Canada), 25 in Europe (19 in the UK, one in Finland, two in Germany, one in Switzerland, and two in Greece), and one in Asia (Malaysia). Around 38.6% ( $n=17$ ) of CPGs did not use any system to evaluate the level of evidence or grading of recommendations, whereas 25 CPGs used different systems, and two CPGs did not declare the system's name. Only eight CPGs declared the use of the AGREE reporting checklist, and 34 did not follow any reporting checklist to develop the guidelines. Although eight guidelines stated that they used the AGREE reporting checklist, our observations revealed that five of them did not adhere to all the items adequately. These areas require further modification

in future updates, particularly regarding the declaration of the guideline development group, transparent systematic search methods, clear statement of the inclusion and exclusion criteria for the evidence, and consideration of the risk, side effects, and benefits in formulating recommendations. Nineteen guidelines declared their funding source, but of those, 10 did not provide information about the role of the funders. On the contrary, 20 CPGs provided a comprehensive conflict of interest statement, including a declaration of how competing interests influenced the guideline process. Two CPGs, however, stated that information about competing interests was available upon request. The years since publication of the guidelines ranged from 0 to 21, and the number of authors varied from 1 to 26.

### 3.3 | Quality assessment of guidelines (AGREE II scores) and regression analysis

Table 2 presents the mean score (95% CI) and minimum and maximum values for all six domains of the AGREE II tool for the included guidelines. The lowest mean score

TABLE 1 Characteristics of included guidelines ( $n = 44$ ).

Characteristics	All guidelines, $n$ (%)
Year of publication	
2000–2006	5 (11.4)
2007–2011	10 (22.7)
2012–2016	13 (29.5)
2017–2021	16 (36.4)
Development organizations	
AAPD	10 (22.7)
EAPD	7 (15.9)
Finnish Medical Association	1 (2.3)
AWMF, DGZ, DGZMK	2 (4.5)
SDCEP	4 (9.1)
WHO	1 (2.3)
Ministry of Health Malaysia	1 (2.3)
USPSTF	1 (2.3)
BSPD	5 (11.4)
UK national guideline	4 (9.1)
NICE	1 (2.3)
ADA	3 (6.8)
IADT	4 (9.1)
Number of centers <sup>a</sup>	
$\leq 2$	36 (81.8)
$\geq 3$	8 (18.2)
Continent of published guidelines	
Europe	25 (56.8)
Others <sup>b</sup>	19 (43.2)
Reporting checklist to conduct guidelines	
No reporting checklist	34 (77.3)
AGREE checklist	8 (18.2)
Other tools <sup>c</sup>	2 (4.5)
Identification of “guideline” or “recommendation” in the title	
No	5 (11.4)
Yes	39 (88.6)
Evidence grading system	
No system	17 (38.6)
Unclear <sup>d</sup>	2 (4.5)
GRADE	12 (27.3)
SIGN	11 (25)
Other grade system <sup>e</sup>	2 (4.5)
Language of published guideline	
English	41 (93.2)
German	2 (4.5)
Finnish	1 (2.3)

TABLE 1 (Continued)

Abbreviations: AAPD, American Academy of Pediatric Dentistry; ADA, American Dental Association; AWMF, DGZ, DGZMK, Association of the Scientific Medical Societies in Germany, German society for tooth preservation, German Society of Dentistry and Oral Medicine; BSPD, British Society of Paediatric Dentistry; EAPD, European Academy of Paediatric Dentistry; IADT, International Association of Dental Traumatology; NICE, National Institute for Health and Care Excellence; SDCEP, Scottish Dental Clinical Effectiveness Programme; USPSTF, US Preventive Services Task Force; WHO, World Health Organization.

<sup>a</sup>When the number of centers is  $\leq 2$ , we considered it single center. Number of centers  $\geq 3$  is considered as multicenter.

<sup>b</sup>Other countries include one from Asia and 18 from Europe.

<sup>c</sup>Other tools such as special tool for the organization (developing NICE guidelines or WHO manual for development of guideline/SIGN handbook).

<sup>d</sup>Unclear as the guideline graded the recommendations and level of the evidence without declaring the system used.

<sup>e</sup>Other systems such as USPSTF.

was for Domain 5: Applicability (16%, 95% CI: 10.8–21.4). The highest mean score was for Domain 4: Clarity of Presentation (58%, 95% CI: 50.8–64.9). Only one guideline conducted by Scottish Dental Clinical Effectiveness Programme (SDCEP), “Conscious Sedation in Dentistry-third edition”, was reported with a score  $>60\%$  for all six domains. The quality evaluation of the 44 guidelines with the mean scores of six domains of the AGREE II tool and the overall assessment are presented in Appendix S4. The linear regression analysis (univariate and multivariate) is reported in Appendix S5.

### 3.3.1 | Scope and purpose

This domain is concerned with evaluating the guidelines' overall objectives, research question, and target population. Only 34.1% ( $n = 15$ ) guidelines had scores  $\geq 60\%$  and 27.3% ( $n = 12$ ) had scores  $\leq 30\%$ . On multivariate analysis, the reporting quality of this domain increased significantly with the use of reporting checklist ( $\beta = .197$ ,  $p = .007$ ), with the authors declared the name of system used to evaluate level of evidence ( $\beta = .155$ ,  $p = .014$ ), and with increase in the number of authors ( $\beta = .014$ ,  $p = .004$ ).

### 3.3.2 | Stakeholder involvement

This domain is related to whether the guideline was developed by declared stakeholders, the involvement of relevant professional groups, and whether the developers have considered the views and preferences of the target population. Seven guidelines were reported with scores

≥60%, whereas 54.5% ( $n=24$ ) guidelines presented with scores ≤30%. On multivariate analysis, it was observed that the reporting quality of this domain significantly improved for guidelines that involved multicenter studies ( $\beta=.218$ ) compared with those based on single-center studies. Additionally, guidelines that utilized a reporting checklist demonstrated a higher overall quality ( $\beta=.181$ ,  $p=.017$ ) than guidelines that did not follow a reporting one.

**TABLE 2** Mean score (95% confidence interval), and minimum and maximum values for all the six domains of the AGREE II tool ( $n=43$ ).

Domain	Mean	Min-Max %	95% CI
D1: Scope and Purpose	48	8–94	40.5–56.2
D2: Stakeholder Involvement	32	0–92	24.5–40.2
D3: Rigour of Development	29	2–83	22.3–35.2
D4: Clarity of Presentation	58	6–100	50.8–64.9
D5: Applicability	16	0–63	10.8–21.4
D6: Editorial Independence	39	0–100	27.1–50.4

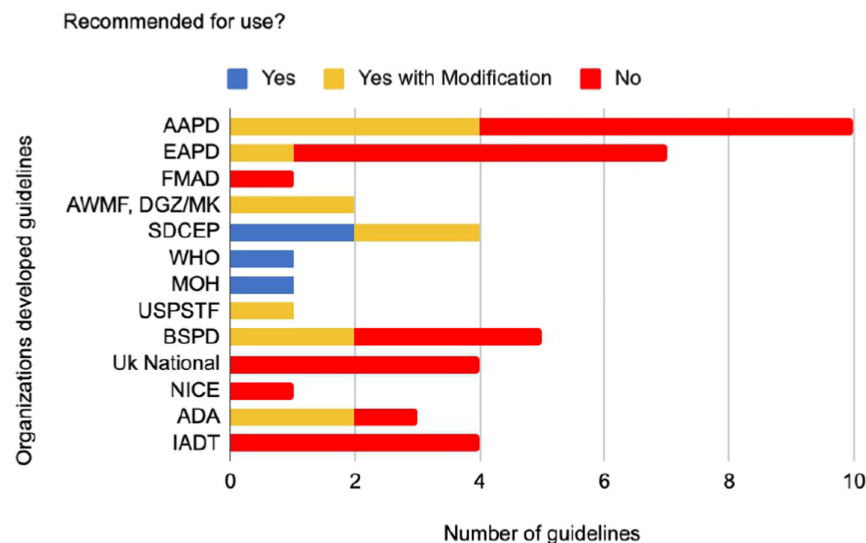
### 3.3.3 | Rigour of development

This domain concerns methodological search, the process of formulating the recommendations, and future update methods. Only 9.1% ( $n=4$ ) of guidelines presented scores ≥60%, whereas the majority ( $n=27$ ) of the guidelines reported scores ≤30%. On multivariate analysis, we found that the reporting quality of this domain experienced a substantial improvement ( $\beta=.203$ ,  $p=.004$ ) when a reporting checklist was used in the development of CPGs. Moreover, the reporting quality of Domain 3 demonstrated an increase ( $\beta=.147$ ,  $p<.001$ ) when the authors declared the name of the system used to evaluate the level of evidence.

### 3.3.4 | Clarity of presentation

This domain addresses the presentation and format of guidelines. Generally, most of the guidelines performed well; 20 CPGs presented scores ≥60%, whereas five guidelines reported scores ≤30%. On multivariate analysis, we identified a significant improvement in this domain: when

The “overall assessment” whether recommended for use, recommended with modification, or not recommended.



**FIGURE 2** Overall assessment of the AGREE II tool (a recommendation for use for each guideline). Depending on the scores of Domains 3, 4, and 5, we considered the guidelines recommended if two domain scores were ≥60%, recommended for use if two domain scores were between 30% and 60%, and not recommended for use if two domain scores were ≤30%. We recommended four guidelines for use, 14 guidelines recommended for use with modifications, and 25 guidelines were not recommended for use. Abbreviations: AAPD, American Academy of Pediatric Dentistry; ADA, American Dental Association; AWMF, DGZ, DGZMK, Association of the Scientific Medical Societies in Germany, German society for tooth preservation, German Society of Dentistry and Oral Medicine; BSPD, British Society of Paediatric Dentistry; EAPD, European Academy of Paediatric Dentistry; FMAD, Finnish Medical Association Duodecim; IADT, International Association of Dental Traumatology; MOH, Ministry of Health Malaysia; NICE, National Institute for Health and Care Excellence; SDCEP, Scottish Dental Clinical Effectiveness Programme; USPSTF, US Preventive Services Task Force; WHO, World Health Organization.

using a reporting checklist ( $\beta = .166$ ), when guideline developers declared the system used for evaluating level of evidence ( $\beta = .199$ ), and when guideline developers used (an unclear [unknown] or declared the name) of the system for assessing level of evidence ( $\beta = .400$ ).

### 3.3.5 | Applicability

This domain focuses on evaluating barriers, facilitators, and implementation of CPGs into practice. Only two guidelines were reported with scores  $\geq 60\%$ , whereas 35 guidelines were reported with scores  $\leq 30\%$ . Approximately 90.9% ( $n = 40$ ) of the guidelines did not report the facilitators and potential barriers, whereas 75% ( $n = 33$ ) did not provide information on cost implications. Additionally, 20.5% ( $n = 9$ ) of the guidelines were unclear in reporting cost information as they did not specify the methods used to obtain such data. The results of our multivariate analysis indicated a significant improvement in the reporting quality of this domain for guidelines that declared the use of a reporting checklist ( $\beta = .0166$ ,  $p = .024$ ). Furthermore, there was a slight increase in the reporting quality when guidelines involved a larger number of authors ( $\beta = .008$ ,  $p = .008$ ).

### 3.3.6 | Editorial independence

This domain evaluates the funding declaration and conflict of interest. Nineteen guidelines did not report the funding source and conflict of interest declaration, which received a score of 0%, whereas five guidelines fully reported both funding source and conflict of interest statement and reported with a domain score of 100%. Using multivariate analysis, we identified a slight decrease in the quality of Domain 6 as the years since publication increased ( $\beta = -.017$ ,  $p = .047$ ). Furthermore, we observed a significant improvement in the quality of reporting when the guidelines incorporated a higher number of authors ( $\beta = .029$ ,  $p < .001$ ).

### 3.3.7 | Overall assessment

We recommend four guidelines for use, 14 for use with modification, whereas 26 are not recommended for use (Figure 2). We conducted univariate logistic regression analysis as reported in Table 3. Additionally, we performed multivariate analysis to gain a comprehensive understanding of the contributions of these characteristics to the overall assessment. Our findings revealed that

**TABLE 3** Logistic regression analysis (univariate and multivariate) between the overall rate assessment and the characteristics of the study.

Predictor variables/category	Overall rate assessment			
	Univariate		Multivariate	
	Odds ratio (SE)	<i>p</i> -Value	Odds ratio (SE)	<i>p</i> -Value
Years since publication <sup>a</sup> (per unit)	0.912 (0.047)	.074	—	—
Country				
Europe	Baseline		—	—
Others <sup>b</sup>	0.908 (0.501)	.861		
Number of centers				
$\leq 2$	Baseline		Baseline	
$\geq 3$	10.14 (8.45)	<b>.005*</b>	14.79 (13.89)	<b>.004*</b>
Used a checklist to conduct CPG				
No	Baseline		Baseline	
Yes	48.43 (54.99)	<b>.001*</b>	21.61 (25.21)	<b>.008*</b>
System for level of evidence				
No	Baseline		Baseline	
Yes	7.04 (4.58)	<b>.003*</b>	7.39 (5.57)	<b>.008*</b>
Unclear	1.64 (2.05)	.691	3.78 (5.03)	.319
Number of authors (per unit)	1.16 (0.05)	<b>.001*</b>	—	—

\*Logistic regression analysis considering  $p < .05$  significant level.

<sup>a</sup> Years since publication = the year of publication – the year of our search strategy.

<sup>b</sup> Other continent (Asia and Europe).



guidelines with multicenter involvement had significantly increased overall assessment (OR=14.79). Moreover, guidelines that declared the name of the system used for the level of evidence demonstrated a significant increase (OR=7.39) in overall assessment. Furthermore, overall assessment was significantly enhanced for guidelines that declared the use of a reporting checklist (OR=21.61,  $p=.008$ ).

## 4 | DISCUSSION

Our study is the most recent and comprehensive evaluation of CPGs in paediatric dentistry. We analyzed 44 paediatric dental guidelines using the AGREE II tool without language or publication year restrictions. Our findings show poor overall quality, with notable flaws in Domains 2, 3, and 5. These shortcomings are concerning as they may introduce bias into the recommendations, mislead readers into unreliable decisions, and neglect the importance of including the views and preferences of the target population. Since these CPGs aim to provide proper health care and improve patient outcomes, it is crucial to use a standardized appraisal instrument to evaluate the reporting quality of these guidelines.

Our results are consistent with those of previous studies in different medical fields, such as CPGs for rehabilitation after anterior cruciate ligament reconstruction<sup>19</sup> and in critical care CPGs.<sup>20</sup> Additionally, in the dental field, quality assessments of CPGs, such as paediatric dentistry,<sup>8</sup> orthodontics,<sup>21</sup> common dental procedures,<sup>22</sup> and dental clinical practice CPGs,<sup>23</sup> have been carried out and the quality of dental guidelines has been found to be poor and inadequate regarding the standards set by the AGREE instrument. Our findings suggest that there is a pressing need to improve the quality of paediatric dental guidelines to ensure that they are reliable, evidence-based, and meet the needs of the target population.

Contrariwise, one study that evaluated the quality of traditional medicine CPGs<sup>24</sup> reported a poor quality for “Clarity of Presentation,” and the key recommendations were not easy to find. Most of our included guidelines and recommendations were well-described, and the key recommendations were well-organized. In our study, the only guideline that obtained a score of 100% for Clarity of Presentation was developed by SDCEP and other collaborators’ institutes (prevention and management of dental caries in children “second edition”). Almost all guidelines neglected to address the “Applicability” domain, as they failed to report on key aspects such as facilitators and potential barriers, cost implications, and methods used to gather cost information. This may lead to guidelines that recommend costly interventions or not-applicable

recommendations for certain healthcare situations. Future guidelines should include the cost implications with an economic analysis and the potential impact of the recommendations on resources. Furthermore, CPGs that have not reported any monitoring or audit criteria might not provide appropriate methods for implementation of the recommendations into practice. We strongly urge developers to provide explicit examples of facilitators, such as ensuring practitioners possess the required skills to implement recommended care, as well as barriers, such as inadequate availability of necessary equipment for specific interventions. Additionally, it is crucial to clearly articulate how the identified facilitators and barriers influenced the guideline development process and shaped the formulation of recommendations. This level of detail will greatly assist practitioners in effectively applying the recommendations within various paediatric dental settings.

Following orthodontic CPGs<sup>21</sup> for the domain “Rigour of Development,” we observed the declaration of the systematic methods or the complete search strategy and eligibility criteria to be lacking. 31.8% ( $n=14$ ) have flaws in systematic methodology, eight guidelines did not have a systematic search, and seven guidelines conducted a systematic review and considered it the base for their recommendations, which may result in potential bias and compromise the validity of the recommendations. The potential bias could be avoided when the developers of the CPGs perform a methodological search and declare the complete search strategy and detailed analysis for the level of evidence through proper tools such as GRADE inside the guideline to ensure transparency. In addition, the “Stakeholder Involvement” domain received a potentially lower score; therefore, we suggest including a wider range of experts and patient representatives and involving them throughout the development of the guideline. Attention should be given to Domain 6, “Editorial Independence,” as certain guidelines did not provide information about the role of funders in the development of the CPG, and some stated that conflict of interest disclosures were available upon request. Our findings indicated a relatively low mean score in this domain, highlighting a lack of transparency that raises concerns regarding potential influences or biases that may have been introduced during the guideline development process. Several studies suggested that increasing the awareness of this domain is mandatory, as it could be considered another source of bias.<sup>20,25,26</sup> It is recommended to declare the funding and conflict of interest statement inside the guideline for the readers to ensure transparency and align to Open Science Practice concepts.

We conducted a regression analysis to examine the factors that could potentially influence the reporting quality of the AGREE II domains. Based on our findings, we can conclude that declaring the use of a reporting

checklist contributes to the improvement of all domains except for Domain 6 (Editorial Independence). It is worth noting that the estimated coefficient for improvement in all five domains suggested a slight improvement. This could be attributed to the fact that developers may have declared the use of the reporting checklist without fully adhering to all the items. The exclusion of the “reporting checklist” characteristic from the multivariate linear regression analysis for Domain 6 (Editorial Independence) can be attributed to the specific focus and characteristics of this domain, which may not be directly influenced by checklist use. Editorial Independence is more concerned with the ethical and governance aspects of guideline development. Future research should explore specific variables to better understand the determinants of Editorial Independence. In relation to Domain 4, our analysis highlighted a significant improvement when the tool used for assessing the level of evidence was declared by name, or it remained unclear. This suggests that the act of declaring the level of included evidence still demonstrates a degree of transparency and allows stakeholders to assess the appropriateness and reliability of the evidence presented. The results from the logistic regression analysis for the overall assessment underscore the importance of multicenter collaboration, adherence to reporting practices, and declared the name of the system used to evaluate the evidence in the guideline development process. Incorporating these practices can enhance the quality and reliability of guidelines, providing valuable insights for guideline developers and users.

The AGREE collaboration defined quality of guideline as the confidence that the potential biases of guideline development have been addressed adequately and that the recommendations are both internally and externally valid and feasible for practice.<sup>27</sup> Furthermore, this tool has been endorsed by the Guideline Review Committee of the World Health Organization (WHO) and recommended for guideline developers and assessors.<sup>28</sup> Our study is in agreement with previous studies,<sup>19–21,23,25,29</sup> by identifying two aspects of the AGREE tool during its application for appraisal purposes. First, in their manual, there is no clear threshold or specific domains that are more significant over others to facilitate the appraisal process and determine the quality of CPGs. Therefore, the quality of the guideline would be an individual appraiser’s decision. We overcame this drawback by searching the literature. We found a systematic review<sup>18</sup> that examined the six domains to determine which domain strongly influences the overall assessment. Second, the AGREE II instrument evaluates the methodological process and guideline structure but does not evaluate the content validity of the recommendations. Moreover, we followed the AGREE II guidelines as closely as possible to ensure that

our assessments were consistent and reliable. Although we did not recommend 25 guidelines for use according to the AGREE tool assessment, they might include good clinical recommendations/advice underpinned by reliable evidence. Nevertheless, as the developer of the guidelines did not use a reporting checklist to formulate it, it resulted in poor reporting methodology.

In our study, two reviewers independently assessed the included guidelines, which increases the reliability of the assessment and strengthens the study findings. We evaluated most of the available paediatric dentistry CPGs, including two in German and one in Finnish. The German CPGs were translated into English using the DeepL translator, and the translations were then reviewed and revised by an expert in the German language. The Finnish CPG was obtained directly in English from the developer to ensure comprehensive evaluation coverage. The NICE organization has decided to close their evidence search in March 2022, which affected the reproducibility of our search strategy for NICE evidence search, but all the included CPGs retrieved from this website are accessible from bibliographic databases “the providers’ websites.” We recognize that many guidelines featured were published before the launch of the AGREE II tool in 2010. Nevertheless, the initial AGREE tool had been validated and accessible as far back as 2001, enabling a retrospective assessment of earlier guidelines.<sup>27</sup> We surmounted this constraint by incorporating solely the latest available version whenever a guideline had been updated. Consequently, we are of the opinion that the advantages of having thorough coverage of paediatric dentistry outweighed the slight risk associated with incorporating only one guideline from the year 2000.

Our findings reveal the need to raise awareness for the future guideline of paediatric dentistry updates to reporting quality for domains “Applicability,” “Rigour of Development,” “Stakeholder Involvement,” and “Editorial Independence.” Correspondingly, we recommend implementing the AGREE reporting checklist while conducting the CPGs, as they contain direct recommendations to paediatric dentistry practitioners. Paediatric dentists should be aware of this while implementing recommendations into practice. Future research would be valuable to assess the reporting quality of the best practice documents using a validated appraisal tool.

#### AUTHOR CONTRIBUTIONS

D.P.R., M.M.B., F.M.M., R.A.E., and C.L.G. were involved in conception and design of the study. R.A.E. and C.M.L. were involved in data acquisition. R.A.E., T.G.C., F.M.M., M.S.C., and D.P.R. were involved in data analysis and interpretation. R.A.E. and C.M.L. drafted the manuscript. D.P.R., M.M.B., F.M.M., M.S.C., T.K.T., T.G.C., and C.L.G. revised and gave final approval of the manuscript.

## ACKNOWLEDGMENTS

We would like to thank Dr. Jonas de Almeida Rodrigues for his kind help in revising the translation of the included German-language guidelines.

## FUNDING INFORMATION

This study was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES; Coordination for the Improvement of Higher Education Personnel)—finance code 001 granting scholarships to R.A.E. and C.L.G. Conselho Nacional de Desenvolvimento Científico e Tecnológico (National Council for Scientific and Technological Development; CNPq) provides scholarship awards for research productivity in Brazil to Drs. Raggio, Braga, and Mendes. The funders had no role in conducting the study, interpreting the data, or deciding to submit the manuscript for publication.

## CONFLICT OF INTEREST STATEMENT

All the authors declare that they have no competing interests.

## DATA AVAILABILITY STATEMENT

All data generated or analyzed during this study are included in Appendix S6.

## REGISTRATION

The protocol of this study was prospectively registered on Open Science Framework—DOI: [10.17605/OSF.IO/BFNGW](https://doi.org/10.17605/OSF.IO/BFNGW).

## ORCID

Rokaia Ahmed Elagami  <https://orcid.org/0000-0002-7238-4598>

Tamara Kerber Tedesco  <https://orcid.org/0000-0003-0794-1578>

Fausto Medeiros Mendes  <https://orcid.org/0000-0003-1711-4103>

## REFERENCES

- American Academy of Pediatric Dentistry. *Overview. The Reference Manual of Pediatric Dentistry*. American Academy of Pediatric Dentistry; 2021:7-9.
- Petersen PE, Bourgeois D, Ogawa H, Estupinan-Day S, Ndiaye C. The global burden of oral diseases and risks to oral health. *Bull World Health Organ*. 2005;83(9):661-669.
- Straus SE, Glasziou P, Richardson WS. *Evidence Based Medicine*. 4th ed. Elsevier; 2011.
- Chiappelli F. Evidence-based dentistry: two decades and beyond. *J Evid Based Dent Pract*. 2019;19(1):7-16. doi:10.1016/j.jebdp.2018.05.001
- Chiappelli F. *Fundamentals of Evidence-Based Health Care and Translational Science*. Springer; 2014.
- Institute of Medicine (US). Committee on clinical practice guidelines. In: Field MJ, Lohr KN, eds. *Guidelines for Clinical Practice: from Development to Use*. National Academies Press (US); 1992.
- Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ*. 1999;318(7182):527-530. doi:10.1136/bmj.318.7182.527
- Shah P, Moles DR, Parekh S, Ashley P, Siddik D. Evaluation of pediatric dentistry guidelines using the AGREE instrument. *Pediatr Dent*. 2011;33(2):120-129.
- Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383-394. doi:10.1016/j.jclinepi.2010.04.026
- Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ*. 2010;182(18):E839-E842. doi:10.1503/cmaj.090449
- Brouwers MC, Spithoff K, Kerkvliet K, et al. Development and validation of a tool to assess the quality of clinical practice guideline recommendations. *JAMA Netw Open*. 2020;3(5):e205535. doi:10.1001/jamanetworkopen.2020.5535
- AGREE Next Steps Consortium. The AGREE II Instrument [Electronic version]. 2017. Accessed March 1, 2023. <http://www.agreetrust.org>
- The AGREE Reporting Checklist: a tool to improve reporting of clinical practice guidelines. *BMJ*. 2016;354:i4852. doi:10.1136/bmj.i4852
- Vernooij RW, Alonso-Coello P, Brouwers M, Martínez García L, CheckUp Panel. Reporting items for updated clinical guidelines: checklist for the reporting of updated guidelines (CheckUp). *PLoS Med*. 2017;14(1):e1002207. doi:10.1371/journal.pmed.1002207
- American Academy of Pediatric Dentistry. Overview. *The Reference Manual of Pediatric Dentistry*. American Academy of Pediatric Dentistry; 2022:7-9.
- Florez ID, Brouwers MC, Kerkvliet K, et al. Assessment of the quality of recommendations from 161 clinical practice guidelines using the Appraisal of Guidelines for Research and Evaluation-Recommendations Excellence (AGREE-REX) instrument shows there is room for improvement. *Implement Sci*. 2020;15(1):79. doi:10.1186/s13012-020-01036-5
- Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev*. 2016;5(1):210. doi:10.1186/s13643-016-0384-4
- Hoffmann-Eßer W, Siering U, Neugebauer EA, Brockhaus AC, Lampert U, Eikermann M. Guideline appraisal with AGREE II: systematic review of the current evidence on how users handle the 2 overall assessments. *PLoS One*. 2017;12(3):e0174831. doi:10.1371/journal.pone.0174831
- Andrade R, Pereira R, van Cingel R, Staal JB, Espregueira-Mendes J. How should clinicians rehabilitate patients after ACL reconstruction? A systematic review of clinical practice guidelines (CPGs) with a focus on quality appraisal (AGREE II). *Br J Sports Med*. 2020;54(9):512-519. doi:10.1136/bjsports-2018-100310
- Chen Z, Hong Y, Liu N, Zhang Z. Quality of critical care clinical practice guidelines: assessment with AGREE II instrument. *J Clin Anesth*. 2018;51:40-47. doi:10.1016/j.jclinane.2018.08.011

21. Tejani T, Mubeen S, Seehra J, Cobourne MT. An exploratory quality assessment of orthodontic clinical guidelines using the AGREE II instrument. *Eur J Orthod.* 2017;39(6):654-659. doi:10.1093/ejo/cjx017
22. Faggion CM Jr. Clinician assessment of guidelines that support common dental procedures. *J Evid Based Dent Pract.* 2008;8(1):1-7. doi:10.1016/j.jebdp.2007.10.001
23. Mubeen S, Patel K, Cunningham Z, et al. Assessing the quality of dental clinical practice guidelines. *J Dent.* 2017;67:102-106. doi:10.1016/j.jdent.2017.10.003
24. Yao L, Chen Y, Wang X, et al. Appraising the quality of clinical practice guidelines in traditional Chinese medicine using AGREE II instrument: a systematic review. *Int J Clin Pract.* 2017;71(5):e12931. doi:10.1111/ijcp.12931
25. Chiappini E, Bortone B, Galli L, de Martino M. Guidelines for the symptomatic management of fever in children: systematic review of the literature and quality appraisal with AGREE II. *BMJ Open.* 2017;7(7):e015404. doi:10.1136/bmjopen-2016-015404
26. Yang C, Zhang Z, Zhang L, et al. Quality assessment of clinical practice guidelines on tic disorders with AGREE II instrument. *Psychiatry Res.* 2018;259:385-391. doi:10.1016/j.psychres.2017.08.060
27. 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. doi:10.1136/qhc.12.1.18
28. World Health Organization. *WHO Handbook for Guideline Development.* 2nd ed. World Health Organization; 2014. <https://apps.who.int/iris/handle/10665/145714>
29. Holmer HK, Ogden LA, Burda BU, Norris SL. Quality of clinical practice guidelines for glycemic control in type 2 diabetes mellitus. *PLoS One.* 2013;8(4):e58625. doi:10.1371/journal.pone.0058625

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Elagami RA, Laux CM, Gallegos CL, et al. Do the clinical practice guidelines for paediatric dentistry meet the quality standards? A meta-research and quality appraisal using the AGREE II tool. *Int J Paediatr Dent.* 2023;00:1-12. doi:[10.1111/ipd.13133](https://doi.org/10.1111/ipd.13133)