The Development of Evidence-Based Guidelines in Dentistry

C. M. Faggion Jr., Dr.med.dent.

Abstract: Use of guidelines is an important means of reducing the gap between research and clinical practice. Sound and unbiased information should be available to enable dental professionals to provide better clinical treatment for their patients. The development of clinical guidelines in dentistry should follow standard and transparent methodology. The purpose of this article is to propose important steps for developing evidence-based clinical recommendations in dentistry. Initially, dental guidelines should be extensively sought and assessed to answer focused clinical questions. If there is a paucity of guidelines or if existing guidelines are not of good methodological quality, systematic reviews should be searched or conducted to serve as a basis for the development of evidence-based guidelines. When systematic reviews are produced, they should be rigorous in order to provide the best evidence possible. In the last phase of the process, the overall quality of evidence should be scrutinized and assessed, together with other factors (balance between treatment effects and side effects, patients' values, and cost-effectiveness of therapy) to determine the strength of recommendations. It is expected this approach will result in the development of sound clinical guidelines and consequent improvement of dental treatment.

Dr. Faggion is Colgate Senior Lecturer, Department of Oral Sciences, Dental School, University of Otago, Dunedin, New Zealand. Direct correspondence and requests for reprints to Dr. Clovis Mariano Faggion Jr., Faculty of Dentistry, University of Otago, PO Box 647, Dunedin 9054, New Zealand; 64-3-479-7036 phone; 64-3-479-0673 fax; clovisfaggion@yahoo.com.

Keywords: evidence-based dentistry, clinical practice, clinical guidelines, clinical education, dental education

Submitted for publication 1/30/12; accepted 7/11/12

Clinical guidelines are considered to be very important in providing information about therapeutic approaches in a language that is accessible to most clinicians, bringing potential benefits to practitioners and patients.[1,2] In other words, clinical guidelines may close the gap between scientific evidence and clinical practice. It is therefore imperative that these recommendations are developed in a systematic and transparent way to minimize the risk of biased clinical decisions.

Clinical guidelines for prevention and treatment of oral conditions have been published, and some have already been endorsed by dental specialty organizations.[3] These guidelines are usually developed by experts in their field, in some cases after a panel meeting for consensus-based decision making.[4] To allow reproduction (replication) and implementation, the steps necessary for the development of guidelines should be clearly reported. Nevertheless, only a few dental organizations make such steps publicly available.[5] In the process of developing guidelines, researchers should consider factors other than simply evidence in support of treatments, such as efficacy/effectiveness balance, side effects, and treatment costs.[6] A comprehensive assessment of variables related to the decision making process is proposed by the GRADE system,[7] which has been used for determining clinical recommendations in dentistry.[7,8]

The development of guidelines should follow phases that begin with conception of the research question and search for evidence and end with determination of the strength of clinical recommendations, i.e., how confident the clinician can be that a procedure will do more good than harm. Tools that support each of these phases[9-11] have been published in many dental articles.[7,8,12-17] The objective of this article is to report some important steps for developing evidence-based guidelines in dentistry. Emphasis will be given to the GRADE system for determining the quality of evidence and strength of clinical recommendations, which are essential steps in evidence-based guideline development. It is hoped that this approach will improve clinical guidelines for a variety of dental procedures.

Steps for Developing Evidence-Based Guidelines

The steps necessary for the development of evidence-based guidelines are reported in this section, from the conception of clinical questions to the determination of recommendations. Figure 1 summarizes the process of the development of evidence-based clinical guidelines.

Define the Topic/s of Interest

The planning and development of guidelines require the combined efforts of experts working in several fields, for example, clinicians, methodologists, and statisticians.[18] Members of the guidelines panel should define relevant clinical questions to be answered. In this context, the PICO approach (participants, intervention, comparison, outcome)[19] seems appropriate for formulating relevant clinical questions because this framework may guide various aspects of the systematic review process—for example, determination of eligibility criteria, search strategy, data collection, and presentation of findings.[20] Attention should be paid to issues considered pivotal when formulating questions, such as which measures of outcome are relevant or not to the patients.[21] Consider the development of guidelines for treatment of periodontitis, in which the research question should first be focused on endpoints such as tooth loss and quality of life, which may better represent the efficacy and/or effectiveness of therapy. Surrogate endpoints—for instance, reduction of pocket probing depth (PPD) and changes of clinical attached level (CAL), although most often used for assessing periodontal therapy—may not always correspond to true endpoints such as tooth loss.[22]

Search for Current Guidelines

To avoid duplication of work and waste of resources, current guidelines should be sought and checked for quality. As a general rule, a good balance between sensitivity and precision should be achieved when seeking guidelines. Sensitivity is related to the number of relevant reports identified divided by the total number of relevant reports in existence, and precision is related to the number of relevant reports identified divided by the total number of reports identified.[23] Increasing sensitivity will imply more time and investment of resources; this should be carefully taken into account by members of the panel assessing the feasibility of performing such a search.

Initially, guidelines should be sought electronically on the websites of regional and international dental specialty organizations. These organizations may make clinical recommendations available in the form of statements or reports that may or may not be published in specialized journals. In a second stage, specific search strategies should be used to find existing guidelines in such electronic databases as Medline, Embase, the Cochrane Database of Systematic Reviews, LILACS, and CINAHL. Searching based on pursuing references of references (called "snowballing")[24] should also be performed to increase sensitivity in the retrieval of publications. Unpublished literature should be also searched to ensure the literature search is comprehensive.[25] It seems reasonable to perform a combination of hand-searching and electronic searching to achieve the most comprehensive approach in identifying studies.[26] There should, furthermore, be no language restriction, to enable the inclusion of guidelines published in languages other than English. All these measures would reduce the risk of publication bias, which is regarded as a source of statistical heterogeneity in meta-analysis.[27] A comprehensive search for guidelines is, therefore, possible only if the members of the panel are able to search and assess evidence in a wide variety of languages.

Assess the Quality of Guidelines

Guidelines should be methodologically sound to provide clinical recommendations with the least bias possible. Some evidence suggests that guidelines in dentistry lack good quality.[28] Assessment of the quality of current guidelines selected by the panel is therefore indispensable.

The Appraisal of Guidelines for Research and Evaluation (AGREE) instrument was developed to assess the methodological rigor of developed guidelines by focusing on the transparency of their development.[29] The first validated version of the AGREE instrument was created in 2003 by a group of guideline developers and researchers who hoped to improve confidence in the internal and external validity of guidelines. The original AGREE instrument had twenty-three items divided into six quality domains.[29] The main idea of using such a tool is to provide standard information about the process of guideline development and reporting of this process in the guideline. A second version of the instrument

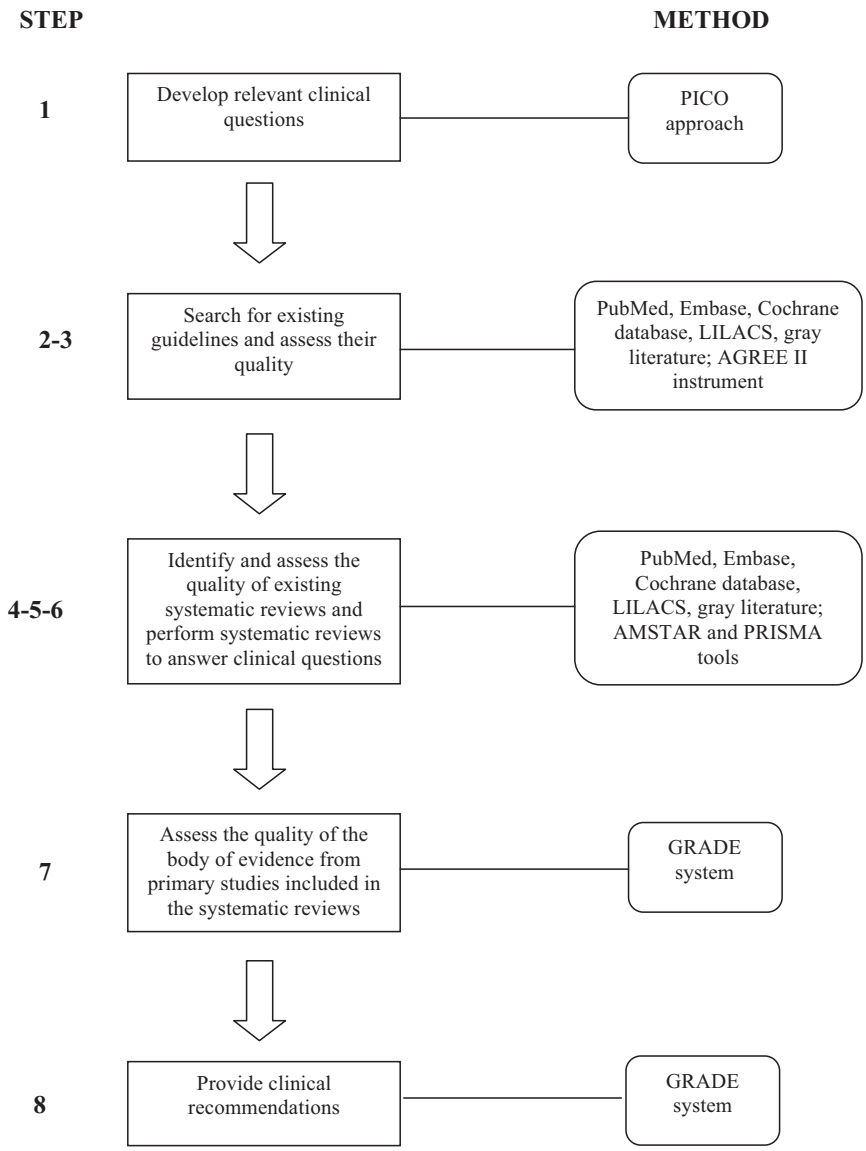

Figure 1. Framework for planning and developing guidelines in dentistry

Note: Steps four to eight involve detailed assessment of the quality of existing guidelines. If a guideline selected in step two is judged to be of good quality, there is no need to proceed to further steps to answer the clinical question previously answered by this guideline.

(AGREE II) was published later; in it, the authors worked to improve the reliability and validity of the original instrument (Table 1).[30] The current checklist has been translated into several languages.

Identify Systematic Reviews

Systematic reviews (SRs) seek to provide high-quality evidence that supports or rejects clinical procedures. SRs are essential to provide the necessary information for the development of guidelines. The SRs identified should cover clinical questions not yet answered by existing guidelines or should be used when existing guidelines are regarded as inadequate for clinical use (determined by use of the AGREE II checklist). The approach used to retrieve SRs should be similar to that used to identify guidelines. Specific literature search strategies should be conducted for

Table 1. AGREE II instrument for assessing the quality of guidelines

AGREE II Item
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.

Source: The Appraisal of Guidelines for Research and Evaluation (AGREE). At: www.agreetrust.org/. Accessed: January 1, 2012.

each prespecified clinical question.[31] Again, as in the process of identification of guidelines, members of the panel should standardize literature search strategies to create a good balance between sensitivity and specificity (Table 2). Existing SRs that may be suitable for answering the clinical questions should furthermore be assessed for quality.

Assess Quality of Systematic Reviews

Several instruments have been proposed for assessing the quality of reporting[32] and of the methodological quality[33] of SRs. The assessment of multiple systematic reviews (AMSTAR) tool consists of eleven items with good face and content validity for assessing the methodological quality of SRs (Table 3).[34] AMSTAR is supported by groups such as the Canadian Agency for Drugs and Technologies in Health and the Cochrane Effective Practice and Organization of Care Group (EPOC),[34] and some authors believe AMSTAR is the best instrument for assessing the methodological quality of SRs.[35] Determination of the methodological quality of SRs depends on the number of checklist items satisfied in the systematic review report. The greater the number of items satisfied, the higher the methodological quality of the SRs.

Conduct Systematic Reviews

If no systematic review of good quality can be identified as able to answer the clinical question, a

Table 2. Proposed literature-search strategy for systematic reviews in the Medline and Cochrane databases via PubMed

Search	Keywords
#5	#3 AND #4
#4	search*[Title/Abstract] OR meta analysis [Publication Type] OR meta-analysis[Title/Abstract] OR meta analysis[MeSH Terms] OR review [Publication Type] OR diagnosis[MeSH Subheading] OR associated[Title/Abstract]
#3	#1 AND #2
#2	"topic 2"
#1	"topic 1"

Source: Montori VM, Wilczynski NL, Morgan D, Haynes RB; Hedges Team. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. BMJ 2005;330(7482):68.

systematic review should be conducted. Members of the panel should focus their efforts on preparing SRs that take into account sound methodological principles; use of checklists for reporting these studies may be required.[32] The preferred reporting items for systematic reviews and meta-analyses, or PRISMA, statement is an update of the QUOROM statement and is currently endorsed by a substantial number of organizations and medical and dental journals. The PRISMA checklist is comprised of twenty-seven items informing the reader how to report the several steps necessary for conducting a systematic review (Table 4). Use of this validated tool is expected to lead to more clarity and transparency of reporting in systematic reviews.[32]

Assess Overall Quality of Evidence

In the process of developing evidence-based clinical guidelines, after selection or development of systematic reviews to answer the relevant clinical questions, the overall quality of evidence should be determined. This means that all evidence derived from each primary study included in the systematic reviews should be thoroughly scrutinized. Assessing domains other than those exclusively related to the internal validity of the studies (i.e., risk of bias) can provide more detailed information, increasing confidence in the level of evidence.

Current systems for assessing the quality of evidence and strength of recommendations present some important limitations.[36] An assessment of the sensibility of six systems (relative to levels of evidence and strengths of recommendations) revealed poor agreement among twelve independent assessors for twelve quality criteria.[36] For example, there was no agreement among the assessors that any of the systems was clear and simple. Other criteria, such as the reproducibility of assessments and suitability of the systems for use by professionals, policymakers, and patients, were considered inappropriate for most systems.[36]

The grading of recommendations assessment, development, and evaluation (GRADE) system uses an explicit, comprehensive, and transparent approach to assess the overall quality of evidence and rate the strength of recommendations.[6] In the GRADE system, the level of evidence is categorized into four levels (high, moderate, low, and very low).[6] First, the level of evidence is rated on the basis of the study design, with randomized controlled trials (RCTs) generating high-quality evidence. Nevertheless, some factors (criteria) can reduce confidence in the quality of evidence, leading to a need to downgrade the overall quality:[37,38]

Internal validity: Studies that have limitations and are likely to be at high risk of bias may reduce confidence in the validity of their results. Limitations of RCTs include failure to conceal allocation, failure to blind, loss from follow-up, failure to consider the intention-to-treat principle, stopping too early for apparent benefit, and selective reporting of outcomes according to the results.[39]

Imprecision: This term refers to a low degree of certainty surrounding an estimate of effect with respect to a specific outcome.[40] Imprecision is normally assessed by checking the 95 percent confidence intervals (CIs) around the difference in effect between the intervention and control groups for each outcome.[41] Generally, if the CIs around the estimates of a treatment effect are not sufficiently narrow, the evidence quality can be downgraded.[41]

Inconsistency: This term refers to the degree of lack of similarity in the effect sizes of different studies forming the whole body of evidence.[40]

Table 3. AMSTAR checklist for the methodological assessment of systematic reviews and meta-analyses

AMSTAR	
1. Was an a priori design provided? The research question and inclusion criteria should be established before the conduct of the review.	• Yes • No • Can't answer • Not applicable
2. Was there duplicate study selection and data extraction? There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.	• Yes • No • Can't answer • Not applicable
3. Was a comprehensive literature search performed? At least two electronic sources should be searched. The report must include years and databases used (e.g., Central, Embase, and Medline). Key words and/or MeSH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study and by reviewing the references in the studies found.	• Yes • No • Can't answer • Not applicable
4. Was the status of publication used as an inclusion criterion? The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language, etc.	• Yes • No • Can't answer • Not applicable
5. Was a list of studies (included and excluded) provided? A list of included and excluded studies should be provided.	• Yes • No • Can't answer • Not applicable
6. Were the characteristics of the included studies provided? In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions, and outcomes. The ranges of characteristics in all the studies analyzed (e.g., age, race, gender, relevant socioeconomic data, disease status, duration, severity, or other diseases) should be reported.	• Yes • No • Can't answer • Not applicable
7. Was the scientific quality of the included studies assessed and documented? A priori methods of assessment should be provided (e.g., for effectiveness studies if the author/s chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria). For other types of studies, alternative items will be relevant.	• Yes • No • Can't answer • Not applicable
8. Was the scientific quality of the included studies used appropriately in formulating conclusions? The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review and explicitly stated in formulating recommendations.	• Yes • No • Can't answer • Not applicable
9. Were the methods used to combine the findings of studies appropriate? For the pooled results, a test should be done to ensure the studies were combinable to assess their homogeneity (e.g., chi-square test for homogeneity, I2). If heterogeneity exists, a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (e.g., is it sensible to combine?).	• Yes • No • Can't answer • Not applicable
10. Was the likelihood of publication bias assessed? An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test).	• Yes • No • Can't answer • Not applicable
11. Was the conflict of interest stated? Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.	• Yes • No • Can't answer • Not applicable

Source: Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.

Table 4. PRISMA checklist for the reporting of systematic reviews and meta-analyses

TITLE		
Title	1	Identify the report as a systematic review, meta-analysis, or both.
ABSTRACT		
Structured summary	2	Provide a structured summary including as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.
INTRODUCTION		
Rationale	3	Describe the rationale for the review in the context of what is already known.
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).
METHODS		
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.
Study selection	9	State the process for selecting studies (e.g., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level) and how this information is to be used in any data synthesis.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2) for each meta-analysis.
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.
RESULTS		
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).
Results of individual studies	20	For all outcomes considered (benefits or harms), present for each study: a) simple summary data for each intervention group, b) effect estimates and confidence intervals, ideally with a forest plot.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15).
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see item 16]).
DISCUSSION		
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., health care providers, users, and policymakers).
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias) and at review level (e.g., incomplete retrieval of identified research, reporting bias).
Conclusions	26	Provide a general interpretation of the results in the context of other evidence and implications for future research.
FUNDING		
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.

*Source:*Moher D, Liberati A, Tetzlaff J, Altman DG; the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(6):e1000097.

When inconsistency is large (as measured by tests of heterogeneity, I^2, or by checking the extent of CI overlap) and unexplained, downgrading the quality of evidence for inconsistency might be appropriate.[42]

Indirectness: This term indicates the extent to which the study population, interventions, and outcome measures are heterogeneous across studies.[43] Indirectness of evidence occurs when therapy interventions are not tested in head-to-head comparisons.[43] The greater the indirectness of evidence is, the greater is the probability that the quality should be downgraded.

Publication bias: This phrase refers to the tendency to publish positive research findings more often than statistically non-significant ones.[44] When the body of evidence is formed by uniformly small studies, particularly those sponsored by industries, the possibility of publication bias should be considered.[45]

In contrast to downgrading the quality of evidence from RCTs, lower-level studies (for example, non-randomized controlled trials) may have their level of evidence (normally low level) upgraded to moderate or high when the following factors (criteria) are present:[46]

Large magnitude of effect: When the weak study design (non-randomized) cannot explain the large benefit of the therapy, we can become more confident in the strength of evidence and upgrade it.

Dose-response gradient: A gradient dose-response may increase our confidence in non-randomized studies, in the case that a minimal amount of a medicament promotes greater therapeutic benefits.

Plausible biases present: In the case of biases or confounders in the study, the demonstrated treatment effect might be reduced. In other words, without these confounders, the findings would probably be more prominent.

The GRADE system advises downgrading the quality of evidence by one or two levels (e.g., from a high to a low level), by considering the severity of the limitations in the existing data. The greater the limitations found in the different domains, the greater will be our confidence in downgrading the level of evidence (Figure 2). For example, when the internal validity domain is assessed, if RCTs forming the body of evidence are at a moderate risk of bias, this would be regarded as a serious limitation, and the overall evidence would therefore be downgraded one level (from high to moderate).[39] Furthermore, if most primary trials providing the information were at high risk of bias, this would be regarded as a very serious limitation, and the overall evidence would therefore be downgraded two levels (from high to low).

GRADE recommends viewing the overall quality of evidence as a continuum rather than as discrete categories. In reality, all assessed domains present minimal to very serious limitations that should be considered when up/downgrading the quality of evidence.[47] Therefore, this assessment is sensitive to the knowledge and interpretation of the reviewers, who might have different perspectives. For instance, a body of evidence presenting point estimates of individual studies showing benefit, but with limitations such as great heterogeneity, may lead one reviewer to downgrade the quality of evidence due to inconsistency. However, another reviewer might not downgrade because all of the estimates show the same positive effect, despite their inconsistency in magnitude, across the studies. The focus of the assessment is on transparency instead of reproducibility of results. However, the rationale used for achieving a specific final score should be stated very clearly for the users of the guidelines.

Providing Clinical Recommendations

One of the main features of the GRADE system is clear differentiation of the quality of evidence from the strength of recommendations.[6] Strength of recommendation is the extent to which one can be confident that adherence to the recommendation will do more good than harm. It is categorized as strong or weak.[48] In other words, the strength of a recommendation is related to the balance between the positive and negative effects of intervention; the greater the proportion of positive effects, the stronger the recommendation is for the proposed therapy. This concept is also valid in the opposite direction, when strong or weak recommendations can be suggested against a proposed therapy.[48]

This dichotomous approach has implications for clinicians, policymakers, and patients.[48] For example, in a strong recommendation, the proposed therapy will likely be accepted by most patients (i.e., only a small proportion of patients will reject the therapy), and policymakers will adopt the therapy as a policy in most situations. In cases of weak recommendations, a trade-off between patients accepting and not accepting the therapy might be expected, and policymakers will need substantial discussion before reaching a decision.[48]

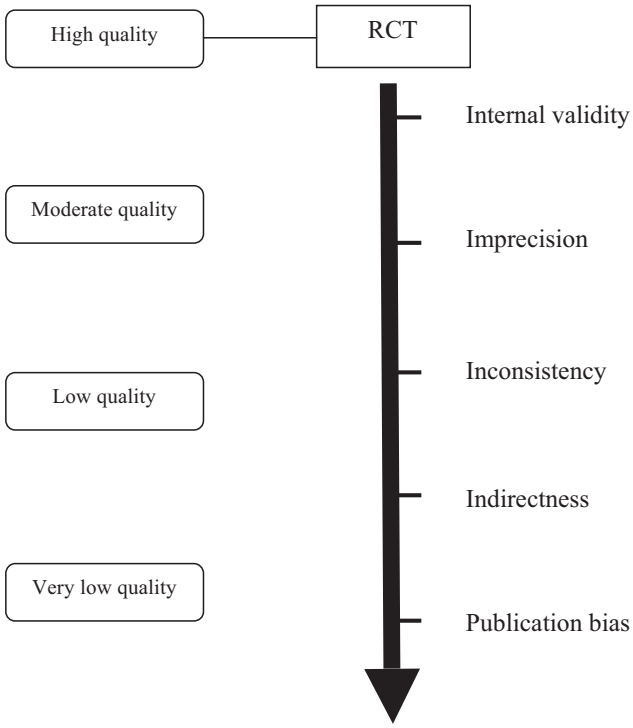

Figure 2. GRADE system for assessment of the quality of the overall body of evidence

Note: The direction of the arrow indicates downgrade of evidence quality when the variables are taken into consideration. The more limitations in the body of evidence, the lower its quality.

Four factors (criteria) are important in determining the strength of a recommendation.[48] First is quality of evidence. In general, a high level of evidence will result in strong recommendations. For example, a body of evidence from RCTs with low risk of bias, without concerns about consistency, precision, and directness of evidence, and low risk of publication bias will generate a strong recommendation. Second is balance between treatment effects and side effects. The greater the uncertainty of the balance between the positive and negative effects of the therapy, the weaker is the confidence in the recommendation. For instance, from the perspective of survival of a single-rooted or multi-rooted tooth with initial furcation involvement in periodontitis cases, open-flap procedures have not been proved to be superior to non-surgical subgingival debridement.[49] Nevertheless, surgical approaches might cause more side effects than non-surgical; therefore, these more invasive procedures might be recommended less strongly than non-surgical periodontal treatment. Third is the values and preferences of the patients receiving the therapy. Patients are all different, and this should be taken into account when guidelines are developed. For example, an older patient is likely to have more concerns about invasive periodontal procedures (for example, guided tissue regeneration) than a younger patient, and this surgical procedure might be weakly recommended for the older patient. Fourth is the cost-effectiveness of therapy. Most patients will consider more expensive therapy with significant but non-clinically relevant results. For example, use of enamel-matrix proteins (EMD) with open flap procedures might be recommended weakly compared with open flap only. EMD has been proved to be more effective than open flap only on the basis of surrogate endpoints, but at higher cost than open flap alone.

Because the GRADE approach assesses different variables for determining the quality of evidence

and the strength of clinical recommendations, a certain level of subjectivity might be expected when performing this assessment. Nevertheless, the emphasis of the GRADE system is on the qualitative assessment of the evidence, to promote more transparent clinical decisions, instead of generating quantitative results for the reproducibility of assessments.

Summary and Conclusions

Figure 2 summarizes the whole concept of the development of evidence-based clinical guidelines. The main objective of this article has been to propose important steps for development of clinical recommendations in dentistry. The use of such an approach may increase dental professionals' confidence in and adherence to guidelines[50] and, consequently, improve clinical treatment. Furthermore, use of a system such as GRADE may improve the level of communication between dental professionals and patients, mainly in cases in which the evidence is low or non-existent. Although currently the concept of a shared decision making process does not seem to be widely used for deciding clinical treatment,[51] participation of patients in the decision making process is becoming an important issue for global discussion.[52] Evidence suggests, moreover, that involvement of the patient in the decision making process may optimize treatment.[53]

Important aspects of the development of guidelines are the accuracy and validation of the several steps performed. Each step, from the identification of existing clinical guidelines to the determination of the overall body of evidence, should be conducted independently and in duplicate by at least two members of the panel, and any disagreements in the assessment should be resolved by discussion. To enable readers to understand the degree of accuracy in assessment of the methodological quality of SRs and guidelines, when possible, a statistical approach should be used to determine the inter-rater reliability of guidelines developers.[54]

In the development of guidelines, SRs are selected to provide primary studies for assessment of the entire body of evidence. To include different study designs that adequately address all relevant outcomes, the selection may include more than one systematic review.[55] Checklists such as AMSTAR will therefore be important in making this process achievable, mainly when several SRs about the same clinical question are found.

Methods for achieving consensus (for example, Delphi) should be used when experts determine the steps in development of guidelines. This method covers a structured process of obtaining information from a group of experts by means of a series of questionnaires.[56] In the Delphi method, experts usually answer the questionnaires anonymously and do not interact directly with each other. This might generate more freedom for making less biased opinions because of the lack of influence of other participants (for example, senior or domineering individuals).[57] Delphi is a valuable method for achieving consensus in the development of guidelines,[57] and it has already been used in the development of such widely accepted documents as some CONSORT checklists.[58] It is not, nevertheless, expected that all members of the guidelines development panel will be able to participate in the guidelines development meetings. Many steps will need to be performed as extra-meeting activities because of the substantial time spent planning and performing these activities. For example, it might not be realistic to plan and develop a systematic review in a short time, normally a three- or five-day meeting. Presence or tele-presence meetings should therefore be reserved for pivotal discussions and decisions in the development of the guidelines.

In conclusion, the methodology proposed is a systematic and transparent approach for developing evidence-based clinical guidelines, taking into account such variables as treatment costs and patients' values. The idea is that leading researchers, clinicians, and policymakers involved in dentistry may follow such an approach to develop recommendations that can improve the treatment of patients.

Acknowledgments

The author would like to thank Dr Nikolaos Nikitas Giannakopoulos for valuable comments that improved this article. Dr. Faggion's position as Colgate Senior Lecturer is partially supported by an unrestricted grant from Colgate-Palmolive New Zealand to the University of Otago.

REFERENCES

1. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ 1999;318(7182):527-30.
2. Ollenschläger G, Marshall C, Qureshi S, Rosenbrand K, Burgers J, Mäkelä M, et al. Improving the quality of health care: using international collaboration to inform guideline programmes by founding the Guidelines International Network (G-I-N). Qual Saf Health Care 2004;13(6):455-60.
3. American Academy of Periodontology. Index of AAP clinical and scientific papers. At: www.perio.org/resources-products/posppr2.html. Accessed: January 1, 2012.

4. Lang NP, Kinane DF, Lindhe J, Sanz M, Tonetti MS. Sixth European workshop on periodontology of the European Academy of Periodontology at the Charterhouse at Ittingen, Thurgau, Switzerland. J Clin Periodontol 2008;35(8 Suppl):1-2.
5. American Dental Association. Clinical recommendations handbook. At: http://ebd.ada.org/contentdocs/FINAL_2011_Revised__ADA_Clinical_Recommendations_Handbook.pdf. Accessed: June 2, 2012.
6. 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.
7. Faggion CM Jr, Giannakopoulos NN, Listl S. How strong is the evidence for the need to restore posterior bounded edentulous spaces in adults? Grading the quality of evidence and the strength of recommendations. J Dent 2011;39(2):108-16.
8. Faggion CM Jr. The shortened dental arch revisited: from evidence to recommendations by the use of the GRADE approach. J Oral Rehabil 2011;38(12):940-9.
9. 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.
10. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. BMJ 2009;339:b2700.
11. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
12. Faggion CM Jr. Clinician assessment of guidelines that support common dental procedures. J Evid Based Dent Pract 2008;8(1):1-7.
13. Ijaz S, Croucher RE, Marinho VC. Systematic reviews of topical fluorides for dental caries: a review of reporting practice. Caries Res 2010;44(6):579-92.
14. Faggion CM Jr, Listl S, Giannakopoulos NN. The methodological quality of systematic reviews of animal studies in dentistry. Vet J, 2011 Sep 14. [Epub ahead of print]
15. Faggion CM Jr, Schmitter M. Using the best available evidence to support clinical decisions in implant dentistry. Int J Oral Maxillofac Implants 2010;25(5):960-9.
16. Suebnukarn S, Ngamboonsirisingh S, Rattanabanlang A. A systematic evaluation of the quality of meta-analyses in endodontics. J Endod 2010;36(4):602-8.
17. Bessa-Nogueira RV, Vasconcelos BC, Niederman R. The methodological quality of systematic reviews comparing temporomandibular joint disorder surgical and non-surgical treatment. BMC Oral Health 2008;26(8):27.
18. Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Clinical guidelines: developing guidelines. BMJ 1999; 318(7183):593-6.
19. Faggion CM Jr, Tu YK. Evidence-based dentistry: a model for clinical practice. J Dent Educ 2007;71(6):825-31.
20. Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions, version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. At: www.cochrane-handbook.org. Accessed: January 30, 2012.
21. Hujoel PP, DeRouen TA. A survey of endpoint characteristics in periodontal clinical trials published 1988-92 and implications for future studies. J Clin Periodontol 1995;22(5):397-407.
22. Hujoel PP, Leroux BG, DeRouen TA, Powell LV, Kiyak HA. Evaluating the validity of probing attachment loss as a surrogate for tooth mortality in a clinical trial on the elderly. J Dent Res 1997;76(4):858-66.
23. Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions. New York: Wiley, 2008.
24. Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ 2005;331(7524):1064-5.
25. Hopewell S, McDonald S, Clarke M, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. Cochrane Database Syst Rev 2007;(2):MR000010.
26. Hopewell S, Clarke M, Lefebvre C, Scherer R. Handsearching versus electronic searching to identify reports of randomized trials. Cochrane Database Syst Rev 2007;(2):MR000001.
27. Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. BMJ 1994;309(6965):1351-5.
28. Faggion CM Jr. Clinician assessment of guidelines that support common dental procedures. J Evid Based Dent Pract 2008;8(1):1-7.
29. 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.
30. 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. CMAJ 2010;182(18):E839-42.
31. Montori VM, Wilczynski NL, Morgan D, Haynes RB; Hedges Team. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. BMJ 2005;330(7482):68.
32. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700.
33. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.
34. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol 2009;62(10): 1013-20.
35. Oxman AD, Schünemann HJ, Fretheim A. Improving the use of research evidence in guideline development: 8. Synthesis and presentation of evidence. Health Res Policy Syst 2006;4:20.
36. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, et al. Systems for grading the quality of evidence and the strength of recommendations I: critical appraisal of existing approaches. BMC Health Serv Res 2004;4(1):38.

37. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64(4):401-6.
38. Faggion CM Jr. Grading the quality of evidence and the strength of recommendations in clinical dentistry: a critical review of 2 prominent approaches. J Evid Based Dent Pract 2010;10(2):78-85.
39. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). J Clin Epidemiol 2011;64(4):407-15.
40. Owens DK, Lohr KN, Atkins D, Treadwell JR, Reston JT, Bass EB, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions—Agency for Healthcare Research and Quality and the effective health care program. J Clin Epidemiol 2010;63(5):513-23.
41. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines: 6. Rating the quality of evidence—imprecision. J Clin Epidemiol 2011;64(12):1283-93.
42. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. J Clin Epidemiol 2011;64(12):1294-302.
43. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. BMJ 2004;328(7454):1490.
44. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. Lancet 1991; 337(8746):867-72.
45. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. J Clin Epidemiol 2011;64(12):1277-82.
46. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al. GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol 2011;64(12):1311-6.
47. Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, et al. GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. J Clin Epidemiol, 2012 Apr 27. [Epub ahead of print]
48. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al. Going from evidence to recommendations. BMJ 2008;336(7652):1049-51.
49. Huynh-Ba G, Kuonen P, Hofer D, Schmid J, Lang NP, Salvi GE. The effect of periodontal therapy on the survival rate and incidence of complications of multirooted teeth with furcation involvement after an observation period of at least 5 years: a systematic review. J Clin Periodontol 2009;36(2):164-76.
50. van der Sanden WJ, Mettes DG, Plasschaert AJ, van't Hof MA, Grol RP, Verdonschot EH. Clinical practice guidelines in dentistry: opinions of dental practitioners on their contribution to the quality of dental care. Qual Saf Health Care 2003;12(2):107-11.
51. Wise J. Shared decision making must move from rhetoric to reality, says King's Fund. BMJ 2011;343:d4734.
52. Salzburg statement on shared decision making. BMJ 2011;342:d1745.
53. Stacey D, Bennett CL, Barry MJ, Col NF, Eden KB, Holmes-Rovner M, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011;(10):CD001431.
54. Gwet KL. Handbook of inter-rater reliability. 2nd ed. Gaithersburg, MD: Advanced Analytics, 2010.
55. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64(4):383-94.
56. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting health care quality indicators: a systematic review. PLoS One 2011;6(6):e20476.
57. Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. PLoS Med 2011;8(1):e1000393.
58. Hopewell S, Clarke M, Moher D, Wager E, Middleton P, Altman DG, Schulz KF; CONSORT Group. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. PLoS Med 2008;5(1):e20.