

APPRAISAL OF GUIDELINES
FOR RESEARCH & EVALUATION
(AGREE)
INSTRUMENT

The AGREE Collaboration

September 2001



AGREE



COPYRIGHT AND REPRODUCTION

This document is the product of an international collaboration. It may be reproduced and used for educational purposes, quality assurance programmes and critical appraisal of clinical practice guidelines. It may not be used for commercial purposes or product marketing. Approved non-English language versions of the AGREE Instrument must be used where available. Offers of assistance in translation into other languages are welcome, provided they conform to the protocol set out by The AGREE Research Trust.

DISCLAIMER

The AGREE Instrument is a generic tool designed primarily to help guideline developers and users assess the methodological quality of clinical practice guidelines. The authors do not take responsibility for the improper use of the AGREE Instrument.

© The AGREE Research Trust, March 2006.
Copyright assigned to The AGREE Research Trust by St George's, University of London (formerly St George's Hospital Medical School)
Published by the AGREE Research Trust

ISBNs 0-9553295-0-7 978-0-9553295-0-0

SUGGESTED CITATION:

The AGREE Collaboration. The Appraisal of Guidelines for Research & Evaluation (AGREE) Instrument, 2001. London: The AGREE Research Trust.

www.agreetrust.org

FUNDING:

The development of the AGREE Instrument was funded by a grant from the EU BIOMED2 Programme (BMH4-98-3669)

FOR FURTHER INFORMATION ABOUT THE INSTRUMENT CONTACT:

Françoise Cluzeau
Email: cluzeau@sgul.ac.uk

or

Jako Burgers
Email: j.burgers@cbo.nl

INTRODUCTION

Purpose of the AGREE Instrument

The purpose of the Appraisal of Guidelines Research & Evaluation (AGREE) Instrument is to provide a framework for assessing the quality of clinical practice guidelines.

Clinical practice guidelines are “systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances”¹. Their purpose is “to make explicit recommendations with a definite intent to influence what clinicians do”².

By quality of clinical practice guidelines we mean the confidence that the potential biases of guideline development have been addressed adequately and that the recommendations are both internally and externally valid, and are feasible for practice. This process involves taking into account the benefits, harms and costs of the recommendations, as well as the practical issues attached to them. Therefore the assessment includes judgements about the methods used for developing the guidelines, the content of the final recommendations, and the factors linked to their uptake.

The AGREE Instrument assesses **both** the quality of the reporting, **and** the quality of some aspects of recommendations. It provides an assessment of the predicted validity of a guideline, that is the likelihood that it will achieve its intended outcome. It does not assess the impact of a guideline on patients’ outcomes.

Most of the criteria contained in the AGREE Instrument are based on theoretical assumptions rather than on empirical evidence. They have been developed through discussions between researchers from several countries who have extensive experience and knowledge of clinical guidelines. Thus the AGREE Instrument should be perceived as reflecting the *current* state of knowledge in the field.

Which guidelines can be appraised with the AGREE Instrument?

The AGREE Instrument is designed to assess guidelines developed by local, regional, national or international groups or affiliated governmental organisations. These include:

1. New guidelines
2. Existing guidelines
3. Updates of existing guidelines

The AGREE Instrument is generic and can be applied to guidelines in any disease area including those for diagnosis, health promotion, treatment or interventions. It is suitable for guidelines presented in paper or electronic format.

¹ Lohr KN, Field MJ. A provisional instrument for assessing clinical practice guidelines. In: Field MJ, Lohr KN (eds). *Guidelines for clinical practice. From development to use*. Washington D.C. National Academy Press, 1992.

² Hayward RSA, Wilson MC, Tunis SR, Bass EB, Guyatt G, for the Evidence-Based Medicine Working Group. Users' guides to the Medical Literature. VIII. How to Use Clinical Practice Guidelines. A. Are the Recommendations Valid? *JAMA*, 1995;**274**, 570-574.

Who can use the AGREE Instrument?

The AGREE Instrument is intended to be used by the following groups:

- i) By *policy makers* to help them decide which guidelines could be recommended for use in practice. In such instances, the instrument should be part of a formal assessment process.
- ii) By *guideline developers* to follow a structured and rigorous development methodology and as a self-assessment tool to ensure that their guidelines are sound.
- iii) By *health care providers* who wish to undertake their own assessment before adopting the recommendations
- iv) By *educators* or *teachers* to help enhance critical appraisal skills amongst health professionals.

Key references

The following sources have been used for developing the AGREE Instrument criteria.

Lohr KN, Field MJ. A provisional instrument for assessing clinical practice guidelines. In: Field MJ, Lohr KN (eds). *Guidelines for clinical practice. From development to use*. Washington D.C. National Academy Press, 1992.

Cluzeau F, Littlejohns P, Grimshaw J, Feder G, Moran S. Development and application of a generic methodology to assess the quality of clinical guidelines. *International Journal for Quality in Health Care* 1999;**11**:21-28.

Grol R, Dalhuijzen J, Mokkink H, Thomas S, Veld C, Rutten G. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. *BMJ* 1998;**317**:858-861.

Lohr KN. The quality of practice guidelines and the quality of health care. In: Guidelines in health care. Report of a WHO Conference. January 1997, Baden-Baden: Nomos Verlagsgesellschaft, 1998.

INSTRUCTIONS FOR USE

*Please read the following instructions carefully
before using the AGREE Instrument*

1. Structure and content of the AGREE Instrument

AGREE consists of 23 key items organised in six domains. Each domain is intended to capture a separate dimension of guideline quality.

Scope and purpose (items 1-3) is concerned with the overall aim of the guideline, the specific clinical questions and the target patient population.

Stakeholder involvement (items 4-7) focuses on the extent to which the guideline represents the views of its intended users.

Rigour of development (items 8-14) relates to the process used to gather and synthesise the evidence, the methods to formulate the recommendations and to update them.

Clarity and presentation (items 15-18) deals with the language and format of the guideline.

Applicability (items 19-21) pertains to the likely organisational, behavioural and costs implications of applying the guideline.

Editorial independence (items 22-23) is concerned with the independence of the recommendations and acknowledgement of possible conflict of interest from the guideline development group.

2. Documentation

Appraisers should attempt to identify all information about the guideline development process prior to appraisal. This information may be contained in the same document as the recommendations or it may be summarised in a separate technical report, in published papers or in policy reports (e.g. guideline programmes). We recommend that you read the guideline and its accompanying documentation fully before you start the appraisal.

3. Number of appraisers

We recommend that each guideline is assessed by at least two appraisers and preferably four as this will increase the reliability of the assessment.

4. Response scale

Each item is rated on a 4-point scale ranging from 4 "Strongly Agree" to 1 "Strongly Disagree", with two mid points: 3 "Agree" and 2 "Disagree". The scale measures the extent to which a criterion (item) has been fulfilled.

- If you are confident that the criterion has been fully met then you should answer "*Strongly Agree*".
- If you are confident that the criterion has not been fulfilled at all or if there is no information available then you should answer "*Strongly Disagree*".

- If you are unsure that a criterion has been fulfilled, for example because the information is unclear or because only some of the recommendations fulfil the criterion, then you should answer "Agree" or "Disagree", depending on the extent to which you think the issue has been addressed.

5. User Guide

We have provided additional information in the User Guide adjacent to each item. This information is intended to help you understand the issues and concepts addressed by the item. Please read this guidance carefully before giving your response.

6. Comments

There is a box for comments next to each item. You should use this box to explain the reasons for your responses. For example, you may "Strongly Disagree" because the information is not available, the item is not applicable, or the methodology described in the information provided is unsatisfactory. Space for further comments is provided at the end of the instrument.

7. Calculating domain scores

Domain scores can be calculated by summing up all the scores of the individual items in a domain and by standardising the total as a percentage of the maximum possible score for that domain.

Example: if four appraisers give the following scores for Domain 1 (Scope & purpose):

	Item1	Item2	Item 3	Total
Appraiser 1	2	3	3	8
Appraiser 2	3	3	4	10
Appraiser 3	2	4	3	9
Appraiser 4	2	3	4	9
Total	9	13	14	36

Maximum possible score = 4 (strongly agree) x 3 (items) x 4 (appraisers) = 48

Minimum possible score = 1 (strongly disagree) x 3 (items) x 4 (appraisers) = 12

The standardised domain score will be:

$$\frac{\text{obtained score} - \text{minimum possible score}}{\text{Maximum possible score} - \text{minimum possible score}} =$$

$$\frac{36 - 12}{48 - 12} = \frac{24}{36} = 0.67 \times 100 = 67\%$$

Note:

The six domain scores are independent and should not be aggregated into a single quality score. Although the domain scores may be useful for comparing guidelines and will inform the decision as to whether or not to use or to recommend a guideline, it is not possible to set thresholds for the domain scores to mark a "good" or "bad" guideline.

8. Overall assessment

A section for overall assessment is included at the end of the instrument. This contains a series of options 'Strongly recommend', 'Recommend (with provisos or alterations)', 'Would not recommend' and 'Unsure'. The overall assessment requires the appraiser to make a judgement as to the quality of the guideline, taking each of the appraisal criteria into account.

SCOPE AND PURPOSE

1. The overall objective(s) of the guideline is(are) specifically described.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

2. The clinical question(s) covered by the guideline is(are) specifically described.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

3. The patients to whom the guideline is meant to apply are specifically described.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

USER GUIDE

SCOPE AND PURPOSE

1. This deals with the potential health impact of a guideline on society and populations of patients. The overall objective(s) of the guideline should be described in detail and the expected health benefits from the guideline should be specific to the clinical problem. For example specific statements would be:

- Preventing (long term) complications of patients with diabetes mellitus;
- Lowering the risk of subsequent vascular events in patients with previous myocardial infarction;
- Rational prescribing of antidepressants in a cost-effective way.

2. A detailed description of the clinical questions covered by the guideline should be provided, particularly for the key recommendations (see item 17). Following the examples provided in question 1:

- How many times a year should the HbA1c be measured in patients with diabetes mellitus?
- What should the daily aspirin dosage for patients with proven acute myocardial infarction be?
- Are selective serotonin reuptake inhibitors (SSRIs) more cost-effective than tricyclic antidepressants (TCAs) in treatment of patients with depression?

3. There should be a clear description of the target population to be covered by a guideline. The age range, sex, clinical description, comorbidity may be provided. For example:

- A guideline on the management of diabetes mellitus only includes patients with non-insulin dependent diabetes mellitus and excludes patients with cardiovascular comorbidity.
- A guideline on the management of depression only includes patients with major depression, according to the DSM-IV criteria, and excludes patients with psychotic symptoms and children.
- A guideline on screening of breast cancer only includes women, aged between 50 and 70 years, with no history of cancer and with no family history of breast cancer.

STAKEHOLDER INVOLVEMENT

4. The guideline development group includes individuals from all the relevant professional groups.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

5. The patients' views and preferences have been sought.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

6. The target users of the guideline are clearly defined.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

7. The guideline has been piloted among target users.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

USER GUIDE

STAKEHOLDER INVOLVEMENT

4. This item refers to the professionals who were involved at some stage of the development process. This may include members of the steering group, the research team involved in selecting and reviewing / rating the evidence and individuals involved in formulating the final recommendations. This item excludes individuals who have externally reviewed the guideline (see Item 13). Information about the composition, discipline and relevant expertise of the guideline development group should be provided.

5. Information about patients' experiences and expectations of health care should inform the development of clinical guidelines. There are various methods for ensuring that patients' perspectives inform guideline development. For example, the development group could involve patients' representatives, information could be obtained from patient interviews, literature reviews of patients' experiences could be considered by the group. There should be evidence that this process has taken place.

6. The target users should be clearly defined in the guideline, so they can immediately determine if the guideline is relevant to them. For example, the target users for a guideline on low back pain may include general practitioners, neurologists, orthopaedic surgeons, rheumatologists and physiotherapists.

7. A guideline should have been pre-tested for further validation amongst its intended end users prior to publication. For example, a guideline may have been piloted in one or several primary care practices or hospitals. This process should be documented.

RIGOUR OF DEVELOPMENT

8. Systematic methods were used to search for evidence.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

9. The criteria for selecting the evidence are clearly described.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

10. The methods used for formulating the recommendations are clearly described.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

11. The health benefits, side effects and risks have been considered in formulating the recommendations.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

USER GUIDE

RIGOUR OF DEVELOPMENT

8. Details of the strategy used to search for evidence should be provided including search terms used, sources consulted and dates of the literature covered. Sources may include electronic databases (e.g. MEDLINE, EMBASE, CINAHL), databases of systematic reviews (e.g. the Cochrane Library, DARE), handsearching journals, reviewing conference proceedings and other guidelines (e.g. the US National Guideline Clearinghouse, the German Guidelines Clearinghouse).

9. Criteria for including / excluding evidence identified by the search should be provided. These criteria should be explicitly described and reasons for including and excluding evidence should be clearly stated. For example, guideline authors may decide to only include evidence from randomised clinical trials and to exclude articles not written in English.

10. There should be a description of the methods used to formulate the recommendations and how final decisions were arrived at. Methods include for example, a voting system, formal consensus techniques (e.g. Delphi, Glaser techniques). Areas of disagreement and methods of resolving them should be specified.

11. The guideline should consider health benefits, side effects, and risks of the recommendations. For example, a guideline on the management of breast cancer may include a discussion on the overall effects on *various* final outcomes. These may include: survival, quality of life, adverse effects, and symptom management or a discussion comparing one treatment option to another. There should be evidence that these issues have been addressed.

12. There is an explicit link between the recommendations and the supporting evidence.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

13. The guideline has been externally reviewed by experts prior to its publication.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

14. A procedure for updating the guideline is provided.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

CLARITY AND PRESENTATION

15. The recommendations are specific and unambiguous.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

USER GUIDE

12. There should be an explicit link between the recommendations and the evidence on which they are based. Each recommendation should be linked with a list of references on which it is based.

13. A guideline should be reviewed externally before it is published. Reviewers should not have been involved in the development group and should include some experts in the clinical area and some methodological experts. Patients' representatives may also be included. A description of the methodology used to conduct the external review should be presented, which may include a list of the reviewers and their affiliation.

14. Guidelines need to reflect current research. There should be a clear statement about the procedure for updating the guideline. For example, a timescale has been given, or a standing panel receives regularly updated literature searches and makes changes as required.

CLARITY AND PRESENTATION

15. A recommendation should provide a concrete and precise description of which management is appropriate in which situation and in what patient group, as permitted by the body of evidence.
- An example of a specific recommendation is: Antibiotics have to be prescribed in children of two years or older with acute otitis media if the complaints last longer than three days or if the complaints increase after the consultation despite adequate treatment with painkillers; in these cases amoxycillin should be given for 7 days (supplied with a dosage scheme).
 - An example of a vague recommendation is: Antibiotics are indicated for cases with an abnormal or complicated course.

However, evidence is not always clear cut and there may be uncertainty about the best management. In this case the uncertainty should be stated in the guideline.

16. The different options for management of the condition are clearly presented.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

17. Key recommendations are easily identifiable

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

18. The guideline is supported with tools for application.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

APPLICABILITY

19. The potential organisational barriers in applying the recommendations have been discussed.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

USER GUIDE

- 16.** A guideline should consider the different possible options for screening, prevention, diagnosis or treatment of the condition it covers. These possible options should be clearly presented in the guideline. For example, a recommendation on the management of depression may contain the following alternatives:
- a. Treatment with TCA
 - b. Treatment with SSRI
 - c. Psychotherapy
 - d. Combination of pharmacological and psychological therapy

- 17.** Users should be able to find the most relevant recommendations easily. These recommendations answer the main clinical questions that have been covered by the guideline. They can be identified in different ways. For example, they can be summarised in a box, typed in bold, underlined or presented as flow charts or algorithms.

- 18.** For a guideline to be effective it needs to be disseminated and implemented with additional materials. These may include for example, a summary document, or a quick reference guide, educational tools, patients' leaflets, computer support, and should be provided with the guideline.

APPLICABILITY

- 19.** Applying the recommendations may require changes in the current organisation of care within a service or a clinic which may be a barrier to using them in daily practice. Organisational changes that may be needed in order to apply the recommendations should be discussed. For example:
- i. A guideline on stroke may recommend that care should be co-ordinated through stroke units and stroke services.
 - ii. A guideline on diabetes in primary care may require that patients are seen and followed up in diabetic clinics.

20. The potential cost implications of applying the recommendations have been considered.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

21. The guideline presents key review criteria for monitoring and/or audit purposes.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

EDITORIAL INDEPENDENCE

22. The guideline is editorially independent from the funding body.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

23. Conflicts of interest of guideline development members have been recorded.

Strongly Agree

4	3	2	1
---	---	---	---

 Strongly Disagree

Comments

USER GUIDE

20. The recommendations may require additional resources in order to be applied. For example, there may be a need for more specialised staff, new equipment, expensive drug treatment. These may have cost implications for health care budgets. There should be a discussion of the potential impact on resources in the guideline.

21. Measuring the adherence to a guideline can enhance its use. This requires clearly defined review criteria that are derived from the key recommendations in the guideline. These should be presented. Examples of review criteria are:

- the HbA1c should be < 8,0%
- the level of diastolic blood pressure should be < 95 mmHg
- If complaints of acute otitis media lasts longer than three days amoxycillin should be prescribed

EDITORIAL INDEPENDENCE

22. Some guidelines are developed with external funding (e.g. Government funding, charity organisations, pharmaceutical companies). Support may be in the form of financial contribution for the whole development, or for parts of it, e.g. printing of the guidelines. There should be an explicit statement that the views or interests of the funding body have not influenced the final recommendations.

Please note: If it is stated that a guideline was developed without external funding, then you should answer 'Strongly Agree'.

23. There are circumstances when members of the development group may have conflicts of interests. For example, this would apply to a member of the development group whose research on the topic covered by the guideline is also funded by a pharmaceutical company. There should be an explicit statement that all group members have declared whether they have any conflict of interest.

FURTHER COMMENTS

Overall Assessment

Would you recommend these guidelines for use in practice?

Strongly recommend

Recommend (with provisos or alterations)

Would not recommend

Unsure