



Assessing the Implementability of Guidelines

**Summary of Report to the
National Institute of Clinical Studies
Prepared for the Guidelines International Network**

October 2006

Acknowledgements

This guide was prepared by Catherine Marshall on behalf of the National Institute of Clinical Studies (NICS).

Copyright

© National Institute of Clinical Studies 2006.

This work is copyright. It may be reproduced in whole or in part for educational use subject to the inclusion and acknowledgement of the source. It may not be reproduced for commercial use or sale.

ISBN: 0-9802934-1-3

Suggested citation: National Institute of Clinical Studies (2006) Assessing the Implementability of Guidelines. NICS, Melbourne.

NATIONAL INSTITUTE OF CLINICAL STUDIES

Fawkner Centre

Level 5 499 St Kilda Rd Melbourne VIC 3004

Phone: 03 8866 0400 Fax: 03 8866 0499

Email: admin@nicssl.com.au

Website: www.nicssl.com.au

ABN 60 095 459 804



Table of contents

Preface	iii
Background.....	1
Discussion.....	2
Summary - Three Level Appraisal	4
Report references.....	5
Appendices	7
Appendix 1 – Project Methods.....	8
Appendix 2- Findings from the literature search.....	11
Appendix 3 - Findings from the informal survey of international guideline developers	16
Appendix 4- Summary of current guideline appraisal tools and instruments.....	26





Preface

Summary of the report on assessing the implementability of guidelines

In July 2006, the National Institute of Clinical Studies (NICS) commissioned a report to describe current international practices for assessing the usability and implementability of guidelines from Catherine Marshall. The report is based on a preliminary review of recent guidelines implementation literature and an informal survey of international guideline developers. The report identifies whether there are suitable, scientifically valid tools or processes for assessing the implementability of guidelines.

The issue of guideline implementability is an issue of concern to all guideline producers and implementers, and is a matter widely discussed within the Guidelines International Network (G-I-N). This summary report is a distillation of the full report prepared for NICS and has been produced as a resource for members of the Guidelines International Network. Hyperlinks are provided throughout this report to allow readers easy access to these tools.

Over the coming months, NICS will be working with both Catherine Marshall and Rick Shiffman to assess the usefulness of the GuideLine Implementability Appraisal instrument (GLIA) for Australian guidelines. A systematic review of the literature relating to the implementability criteria for guidelines will also be undertaken. We plan to publish the findings and will ensure that they are made available to members of the Network.

Finally, we would like to record our thanks to all the people who participated in the survey and shared information and ideas about guideline implementability with us.

Dr Heather Buchan

Chief Executive
National Institute of Clinical Studies
Member, G-I-N Board

Catherine Marshall

Independent Guideline Adviser
Hon Patron, G-I-N
Member, G-I-N Board

October 2006



Assessing the Implementability of Guidelines

Summary Report

Background

Evidence-based guidelines have been developed to improve consumer or patient healthcare outcomes. Many guidelines are now funded by central or regional government agencies seeking to provide the most effective care for consumers within their area of responsibility.

Guidelines are primarily commissioned on topics where there is a significant evidence-practice gap, for example publication of research on the effectiveness of a new medication, equipment, or service design that is shown to be more effective than current practice. Some guidelines also use new research evidence to address uneven patterns of care or unequal access to service by highlighting the groups of consumers who have the greatest (or least) ability to benefit from a new form or configuration of care.

In the last ten years there has been considerable attention placed on improving the quality of the guidelines to ensure that they provide robust, reliable, independent advice. A number of tools, instruments and handbooks have been developed to appraise and describe the rigour required for the development of a “quality” guideline.

Despite the improvements in guideline design, both guidelines research and practical experience show it is possible to develop a guideline that is methodologically sound, but cannot be easily put into practice or implemented as part of a day-to-day healthcare encounter. Health researchers, policymakers and funders, healthcare workers and consumers all find this both frustrating and challenging.

The National Institute of Clinical Studies (NICS) is Australia's national agency for improving health care by helping close important gaps between best available evidence and current clinical practice. NICS helps Australian health care workers and consumers bridge the gap between the research findings and day-to-day clinical practice. As part of its current work plan NICS is aiming to identify ways in which guidelines can be assessed with respect to their ease of implementation. NICS will then be consulting with guideline development organisations to determine whether there are ways that some of these factors might be better incorporated in to guideline development processes in Australia.

This report sets out to inform this work by summarising the latest international experience and advice on assessing the implementability of guidelines. It is based on both a literature search of current guideline implementation research and an informal survey of international guideline developers. The literature search and interviews focused on suitable, scientifically valid guideline assessment tools that would assess the implementability of a guideline. Additional questions were also asked to find out how guideline quality (including usability and implementability) of guidelines is assessed in other countries.

Discussion

Since the late 1990s guideline researchers have demonstrated that many published guidelines fall short of basic quality criteria.^(1, 2) As a result, researchers and guideline developers began identifying the quality criteria/ dimensions that provide an assurance that the guideline has been developed in a methodologically robust manner. These include:

- ensuring the reporting of the scientific findings are accurate and the risk of biased interpretation of the evidence and recommendations is reduced. Guideline assessment tools ask for information about the definition of the topic, search strategies, descriptions of the methodology used to synthesize the evidence and form recommendations, identification of the risks and benefits of treatment or care options;
- reducing the risk of bias of the writers. Assessment tools require that the guideline funding sources are disclosed, multidisciplinary input into guidelines is encouraged and questions are asked about the level of consumer participation and involvement in the formulation of the recommendations. Affiliations of the guideline team members are also recorded, and the names of the individuals who have reviewed the guideline draft are recorded; and
- identifying topics/chapters that should be considered within the guideline. Assessment tools review whether the guideline includes sections on the applicability of the guideline to the local context, the cost of implementing the recommendations, audit and monitoring criteria, dates for updating the guideline.

These dimensions have been incorporated into a number of guideline appraisal instruments. Currently, the [AGREE \(Appraisal of Guidelines for Research and Evaluation Instrument\)](#)⁽¹⁾, and a US instrument produced by a group of guideline developers and implementers [COGS \(Conference on Guideline Standardization\)](#)⁽³⁾ are the most widely used guideline appraisal instrument. Both instruments cover the topics listed above, but have different formats. In Australia, the [National Health and Medical Research Centre \(NHMRC\)](#) does not use AGREE or COGS but has developed [checklists](#) of topics to be included in guideline in order for a guideline to be eligible for NHMRC endorsement.^(4, 5)

Recent research on guideline appraisal acknowledges that guideline appraisal does not address the clinical content of the recommendations nor the quality of the supporting evidence.⁽⁶⁾ In addition, researchers⁽⁷⁾ found that even guidelines explicitly based on evidence frequently fail to address issues of barriers to implementation, monitoring criteria and evidence of pilot testing.

Tools such as AGREE and COGS have a place in assuring that the methodological baseline is robust, and at a very general level assessing the implementability of a guideline. They do not guarantee that a guideline is appropriate for the context of care at either a national/ regional or local level.

To address these issues, a new guideline implementability tool [GuideLine Implementability Appraisal \(GLIA\)](#) was launched in 2005.⁽⁸⁾ It was designed for use prior to publication of the final guideline to assess the implementability of individual recommendations in a guideline. Guideline developers interviewed believe that this tool has considerable promise. This tool is currently used in the US and [NICE](#) in the UK is beginning a trial of the [GLIA](#) instrument over the next two or three months. Full testing and validation of the tool is not yet complete. Also in development is a Handbook for the Adaptation of Guidelines.⁽⁹⁾ This draft handbook recommends that guidelines for adaptation be reviewed using the rigour domain of the AGREE instrument and then outlines other critical issues for consideration of the local

context (including legislation, policy, resource, health care setting, access to equipment and expertise).

Agencies contacted as part of the survey of guideline developers confirmed the importance of consideration of these issues. They described the additional actions and processes they take to assess the “implementability and applicability” of the guidelines (beyond the use of any single appraisal instrument). They all reported that guideline commissioners, developers and endorsers want to know whether a guideline can be applied in the regional / national setting, as well as whether it can be practically implemented at the service delivery level.

Issues that need to be addressed by policy makers and funders include:

- resources (e.g. can we afford it? can we afford not to do it? do we have the equipment needed to make it happen?);
- service configuration (including availability of services, facilities and staff);
- policy and legislation parameters that might have to be modified (including the kinds of incentives that might be put in place to make the guideline easy to use), and whether the guideline is consistent with current policy dimensions, (e.g., does it recognise issues of equality and diversity and cultural/ethnic issues, access for people in regionally remote areas or people with special needs, whether medications are subsidised, whether high cost equipment is widely accessible/ available;
- stakeholder support (e.g., have key stakeholders bought-in to the guideline and do they think it will work?);
- performance Indicators (e.g. what are the health outcomes expected from the change in practice and are clear outcome measures and performance criteria described? How will you know when the guideline recommendations are working?).

A number of checklists and tools for assessing whether guidelines meet these criteria have been developed by guideline development agencies or their central government or health authority. e.g., some tools and checklists address cost impact assessment processes, while others cover equality and diversity. At present there is no single tool that can identify the dimensions or criteria for considering whether a guideline fits within the policy context.

Beyond these high level policy considerations, guideline developers and endorsers want to know about the usability of the guideline in the hands of a person applying the recommendations. This local level assessment/appraisal includes consideration of:

- whether the recommendations are clear about what has to be done, by whom, where, when and how;
- whether local opinion leaders think applying the guidelines would lead to improved health outcomes for consumers;
- whether the local practitioners have sufficient skills and access to resources to apply the guidelines;
- whether there are supporting tools that can be readily used by health practitioners and consumers that will assist them to apply the recommendations (e.g., consumer resources, summary cards) ; and
- whether the guideline information can be easily be translated into computerised formats.

Some of the international guideline developers have designed questionnaires and consultation processes to seek feedback/input on the practical local level of workability of their guidelines and copies of these tools are included either as hyperlinks throughout this report or in Appendix 4.

Summary - Three Level Appraisal

The findings from the literature review and the informal survey of international guideline developers suggest that there is currently no one single tool or process that can determine whether a guideline is methodologically robust, meets national and regional policy and administrative contexts and can be practically applied by local health care workers and consumers. It is also clear that each of these areas of interest must be considered and validated if a guideline is to be put into use.

A three level assessment process to identify the implementability of guidelines is proposed below:

Level 1

The methodological integrity of the guideline should be assessed to ensure that it is scientifically robust and valid. This first step could include an appraisal of the guidelines using the rigour domain of the [AGREE instrument](#). The AGREE instrument is currently the best available, validated tool for assessing the quality of a guideline.

Level 2

The national and regional applicability should be assessed by policy makers, funders and national stakeholder groups to ensure that the guideline could work, that it meets policy standards, that the health benefits accruing from its implementation provide a worthwhile investment and can be supported by the health system. Consideration would include cost impact assessment, availability of services, equipment, staff and medications, clear description of audit and monitoring indicators and other issues such as cultural and geographical appropriateness.

Level 3

Local users “road test” the guideline to see if it can be easily used and ensure that local barriers that could impede its adoption are considered and resolved. This could lead to modification of the guideline. Stakeholder groups could then be invited to endorse the final guideline “products” e.g. summaries, implementation tools, consumer resources. [GLIA](#) appears to be the most useful assessment tool for this level of assessment at present.

Report references

1. Agree Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Quality & Safety in Health Care*. 2003 Feb;12(1):18-23.
2. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. *JAMA*. 1999 May 26;281(20):1900-5.
3. [Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. *Ann Intern Med*. 2003 Sep 16;139\(6\):493-8.](#)
4. [National Health and Medical Research Council. Minimum requirements for formulating NHMRC evidence-based guidelines. NHMRC; 2005.](#)
5. [NHMRC. NHMRC standards and procedures for externally developed guidelines: National Health and Medical Research Council; 2005.](#)
6. Burgers JS. Guideline quality and guideline content: are they related? *Clin Chem*. 2006 Jan;52(1):3-4.
7. Harpole LH, Kelley MJ, Schreiber G, Toloza EM, Kolimaga J, McCrory DC. Assessment of the scope and quality of clinical practice guidelines in lung cancer. *Chest*. 2003 Jan;123(1 Suppl):7S-20S.
8. [Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, et al. The GuideLine Implementability Appraisal \(GLIA\): development of an instrument to identify obstacles to guideline implementation. *BMC Med Inform Decis Mak*. 2005;5:23](#)
9. ADAPTE group. Manual for guideline adaptation [draft]. 2006.
10. Shiffman R, Codish S. The PPI Toolkit and eGLIA.





Appendices

Appendix 1 – Project Methods

Methods

The research for this report was conducted in two parts. A search of current guideline implementation research was undertaken. Secondly, an informal survey of international guideline developers was undertaken to find out:

- how they assessed the quality (including usability and implementability) of guidelines, and;
- if they had developed new tools or approaches that had not been published.

Data sources

Search strategy

A preliminary literature search of the Ovid Medline and Cinahl databases was conducted on July 27 2006. This was based on the search performed by Vlayen et al⁽¹⁾ in their systematic review of clinical practice guideline appraisal tools.

Variations of the following MeSH and text terms were used in combination: practice guidelines, appraisal, and evaluation. To identify papers published subsequent to the search conducted by Vlayen et al⁽¹⁾ results were restricted to those published between 2003 and 2006. A manual search of the references for other relevant articles was also conducted. A number of papers were also provided by the people contacted as part of the informal survey. Due to time restrictions, the search did not include non-English language papers.

A review of the NICS library database was conducted to source articles relating to tools to test the implementability of guidelines.

Results

The Medline search yielded 2230 articles, and the Cinahl search 590 articles, resulting in a total of 2446 articles when these references were combined and duplicates removed. Of these 52 articles were selected. A further 112 articles were identified through the manual searching of references and other specified methods of identifying articles. Because of the need to complete this project within a specified period of time, a preliminary sorting to identify the most important papers was undertaken. Approximately 60 papers were then thoroughly reviewed and the other papers were briefly scanned for relevance. A comprehensive systematic review on implementability is planned for the second phase of this project.

Survey of Guideline Developers

An informal survey of guideline developers was conducted to establish which guideline appraisal instruments were used to assess the implementability and applicability of guidelines. The agencies involved were from the Netherlands USA, Canada, England, Scotland, New Zealand, France, Australia and Germany. Those contacted were:

- Françoise Cluzeau and Ian Saunders, National Institute for Health and Clinical Excellence ([NICE](#));

- Sara Twaddle, Scottish Intercollegiate Guidelines Network ([SIGN](#));
- Leonie Brunt and Steve Caldwell, New Zealand Guidelines Group ([NZGG](#));
- Dave Davis, Guidelines Advisory Committee, Ontario, Canada ([GAC](#));
- Jean Slutsky, Agency for Healthcare Research and Quality ([AHRQ](#));
- Melissa Brouwers, Cancer Care Ontario ([CCO](#));
- Guenter Ollenschlaeger and Monika Lelgemann, Agency for Quality in Medicine, Germany ([AEZQ](#));
- Rick Shiffman, Yale University and member of the American Academy of Pediatrics ([AAP](#));
- Kitty Rosenbrand and Jako Burgers from the Dutch Institute for Healthcare ([CBO](#)), The Netherlands;
- Najoua Mlika-Cabanne from Haute Autorité de Santé ([HAS](#)), France; and
- Cathy Clutton, National Health and Medical Council ([NHMRC](#)) Australia

Note:

Françoise, Jako, Jean and Guenter were involved in the development of the [AGREE Appraisal Instrument](#).

Jean and Rick were involved in the development of the [COGS](#) checklist. Rick is the main developer of the [GLIA](#) instrument.

Jako, Melissa and Najoua are involved in the development of the ADAPTE handbook.

Guenter developed the [DELBI](#) tool.

As a result of the interviews, additional tools and checklist used for the appraisal of guidelines and additional articles were made available to the researcher.

Informal Survey Questions

Each guideline development agency was sent the following list of questions which were used as the basis for the telephone discussion with the researcher.

1. Describe the process your agency uses to assess the quality of clinical practice guidelines. Please indicate when this assessment takes place (eg at draft or pre-publication phases and any groups or trained individuals who are involved at each stage)
2. Does your agency use any of the following appraisal instruments?

AGREE	http://www.agreetrust.org/	Y/N
SHANEYFELT	http://www.openclinical.org/prj_cogs.html	Y/N
GLIA	http://www.biomedcentral.com/1472-6947/5/23	Y/N
DISCERN	http://www.discern.org.uk/discern_instrument.php	Y/N

If another instrument is used which one do you use?

3. Do you have any comments or views about the reasons you don't use the other review instruments?
4. If you use AGREE, do you have a cut off point below which a guideline is not acceptable? If a guideline receives a low score, what action is taken?
5. What level of training is required before the assessors can apply these instruments?
6. How many people assess the same guideline?
7. Do you ask any additional questions or information in addition to using an appraisal instrument, if so what are they? (If you don't use an appraisal instrument, what questions do you ask to determine the quality of a guideline).

8. How do you assess whether a guideline can be implemented?
Please include comment on how you assess:
 - whether the population described for eligibility match the population to which the recommendation is targeted in the local setting;
 - how the intervention meets patient views and preferences in the context of use;
 - if recommendations are compatible with the culture and values in the setting where it is to be used;
 - if the intervention/equipment available in the context of use;
 - if there is the necessary expertise (knowledge and skills) available in the context of use;
 - if there any constraints, organisational barriers, legislation, policies, and resources in the healthcare setting of use that would impede the implementation of the recommendation;
 - the clarity of the recommendations – i.e. Are they unambiguous with specific actions identified?;
 - the complexity or cost involved in translation into electronic/computer format required to implement the guideline; and
 - whether the guideline suggests measurable outcomes and criteria for monitoring and auditing guideline uptake.
9. Does your organisation believe there is a need for a new or modified guideline implementability appraisal tool that looks at the applicability and format of a guideline?
10. If YES, what specific domains or areas would need to be covered by a new appraisal instrument?

References

1. Vlayen J, Aertgeerts B, Hannes K, Sermeus W, Ramaekers D. A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit. *International Journal for Quality in Health Care*. 2005 Jun;17(3):235-42.

Appendix 2- Findings from the literature search

Guidelines are expected to facilitate more consistent, effective and efficient medical practice and improve health outcomes.⁽¹⁾

Effectiveness in daily practice, therefore depends on the guidelines being acceptable and credible.⁽²⁾ However, many published guidelines fall short of basic quality criteria.^(3, 4)

The 2005 [systematic review of appraisal tools for clinical practice guidelines](#)⁽⁵⁾ and the [2000 Comparison of Clinical Practice Guideline Appraisal Instruments](#)⁽⁶⁾ were the starting points for this project. The review by Graham and colleagues assessed 13 guideline appraisal instruments and found there was insufficient evidence to support the use of one single appraisal instrument, although the instrument developed by Cluzeau and colleagues was the most well developed. A new instrument based on the Cluzeau instrument, the [AGREE Instrument](#) was launched.⁽⁷⁾

By 2003, the AGREE instrument had been applied in 11 countries by a large number of guideline appraisers.⁽³⁾ The AGREE instrument is rapidly becoming regarded as the gold standard for guideline appraisal.⁽⁸⁾

The 2005 systematic review by Vlayen and colleagues⁽⁵⁾ set out to identify a tool for appraising guidelines that could serve as a basis for the development of an appraisal tool for clinical pathways. They observed that the AGREE instrument is a validated, easy-to-use, and transparent instrument, which was internationally developed and widely accepted. They also noted some limitation with the AGREE. Firstly, the domain scores are useful for comparing clinical practice guidelines, but it is not possible to set thresholds for the scores to classify a clinical practice guideline as 'good' or 'bad'. Secondly, the AGREE instrument does not assess the clinical content of the clinical practice guideline nor the quality of evidence supporting the recommendations, which is a common deficit in all the existing appraisal tools. These limitations have been acknowledged by one of the AGREE developers.⁽⁹⁾

Similar concerns were also raised by Harpole et al, in an assessment of international guidelines on lung conditions.⁽¹⁰⁾ Fifty-one guidelines were appraised using the AGREE instrument. The reviewers found that even those guidelines explicitly based on evidence, failed to address issues of barriers to implementation, monitoring criteria and evidence of pilot testing. They suggest that well developed guidelines should include consideration of potential barriers to guideline implementation, supply monitoring criteria to assess the guideline's impact and should also provide evidence of pilot testing to ensure that the guideline can be practically put to clinical use.

Both the literature review and survey results identified a number of factors that made the AGREE instrument insufficient and there was general agreement that scientific evidence is only one of the many factors that may influence the translation of research findings to the context of use. The process of considered judgment is essential to guideline development and often requires extensive discussions and consensus among experts.⁽²⁾

Considered judgment has become a formalised process in a number of guideline development agencies around the world. SIGN developed a form that is used by guideline development teams making guideline recommendations to consider:

- the volume of evidence;
- the applicability of the evidence to the local health system;
- generalisability and consistency of the evidence; and

- clinical impact of the proposed intervention ie size of population affected, magnitude of effect, relative benefit over other options, resource implications, balance of risk and benefit.⁽¹¹⁾

CBO has conducted a formal process based on considered judgements to explore other considerations apart from scientific evidence relevant in formulating guideline recommendations.⁽¹²⁾ They see this as a process that can be conducted in parallel with the AGREE instrument.

NICE guidelines and guidance is developed according to a detailed technical manual⁽¹³⁾ setting out the requirements for the preparation of high quality guidelines. This manual is based on the AGREE instrument quality and all guidelines developed under the auspices of NICE must follow the handbook. Guidelines developed outside of the NICE and its Collaborating Centres are routinely assessed using the AGREE instrument.

Applying the handbook requirements has not always resulted in the development of clear, unambiguous advice and recommendations. A 2004 audit of the uptake of guidelines produced by NICE found mixed implementation results. E.g., prescribing of drugs for Alzheimer's disease and prophylactic extraction of wisdom teeth showed trends consistent with, but not obviously a consequence of, the guidance. Prescribing practice often did not accord with the details of the guidance. No change was apparent in the use of hearing aids, hip prostheses, implantable cardioverter defibrillators, laparoscopic hernia repair, and laparoscopic colorectal cancer surgery after NICE guidance had been issued.⁽¹⁴⁾ The auditors concluded that guidance seems more likely to be adopted when:

- there is strong professional support;
- there is a stable and convincing evidence base;
- there is no increased or unfunded costs;
- the organisations implementing the guidance have established good systems for tracking guidance implementation; and
- the professionals involved are not isolated.

In the Netherlands in 2005, a review of 17 pediatric guidelines was conducted using the AGREE instrument.⁽¹⁵⁾ The study found that the AGREE instrument was a useful tool to select high quality guidelines. However, they also acknowledged that guidelines would require adaptation to culture-specific values and local or national pediatric practice.

Lavis et al⁽¹⁶⁾ describe the ways research is used to inform public policy making. They suggest that public policy makers seek research findings that find the best solutions in terms of effectiveness and cost-effectiveness. They also want advice on feasibility (ie.do the research findings reflect the jurisdictional authority, administrative capacities and financial discretion of the health system?). Other elements include acceptability to key stakeholder groups (including consumer, professional associations, government agencies, businesses, donors and international agencies. Lavis et al propose three questions to assist public policymakers assess the local applicability of systematic reviews:

- Could it work?
- Will it work? (or what would it take to make it work?) and;
- Is it worth it?

This analysis reflects many of the ongoing comments in the literature about the questions to be asked of guidelines.

Woolf observes that in an era of limited health care resources, guideline developers must often consider the cost effectiveness or cost utility of interventions. Other policy

considerations he highlights include access to care, availability of qualified personnel and technology, insurance policies, and medicolegal implications.⁽¹⁷⁾ He goes on to observe that guidelines can have adverse policy implications for society if they increase the costs of care, decrease equity, or divert resources from more effective health care interventions. This view is supported by Grimshaw et al.⁽¹⁸⁾ They advise that ideally, the decision of whether and how to develop, disseminate and implement a guideline should be based upon careful estimates of the costs and benefits of the dissemination and implementation strategy, and the costs and benefits of the resulting changes in patient care.

Ouimet and colleagues⁽¹⁹⁾ have sought to identify the factors that induce healthcare decision makers to use clinical guidelines. In the context of resource constraints, 'politicians will not want to pay for services deemed inappropriate by a guideline based on extensive and expensive research. In effect, clinical guidelines do not solely influence "clinical decisions at the bedside", but also the 'rules of operation at hospitals and clinics, and health spending by governments and insurers. Woolf et al defined the potential benefits of clinical guidelines for decision-makers as follows:

Healthcare systems that provide services, and government bodies and private insurers that pay for them, have found that clinical guidelines may be effective in improving efficiency (often by standardizing care) and optimising value for money. Implementation of certain guidelines reduces outlays for hospitalisation, prescription drugs, surgery, and other procedures. Publicising adherence to guidelines may also improve public image, sending messages of commitment to excellence and quality. Such messages can promote good will, political support, and (in some health care systems), revenue.⁽¹⁾

Carolyn Clancy and Kelly Cronin reinforce the importance of buy-in from both policy makers and consumers. The application of science in policy and clinical decision making, however, depends on local or national professional norms, consumers' expectations, and the information infrastructure and tools at the point of care. Applications that support the consumer's role in decision making are a strong focus in the United States. Clinicians must tailor scientific information derived from population-based studies to individual patients' needs and preferences, and policymakers must identify which approaches are most likely to succeed for their programs. Policy interventions, such as the use of formularies, coverage, cost sharing, financial and other incentives have an important albeit indirect impact on clinical decisions.⁽²⁰⁾

Local cultural issues and health service configuration are also recognised factors for the overall usefulness of a guideline. A comparison of four European guidelines on uncomplicated cystitis concluded that there were substantial differences even between high-quality guidelines on the same well defined clinical entity. The selection of literature data and diagnostic and therapeutic recommendations seemed to be influenced by such cultural issues as habits, the patient's expectations and the structure of the healthcare system.⁽²¹⁾

The World Health Organisation (WHO) makes it clear to WHO guideline developers that tradeoffs between the cost of applying the possible recommendations on a population basis, and the population health impacts should be spelt out. In its **Guidelines for WHO Guideline**, WHO advocates that a range of scenarios (or optimal recommendations) should be proposed to cover differing contexts of care, (e.g. where there are very limited resources, or alternatively, where there are unlimited resources). This allows decision makers in different countries to make recommendations based on their process.⁽²²⁾

In addition, guidelines that describe ideal or gold standards for care (even if the resources are not currently available) is seen as useful as long as there is also advice about actions to take in the short and medium term. This approach has been supported by a 2003 Global Consensus Conference that considered breast healthcare guidelines for countries with

limited resources.⁽²³⁾ The summit fully endorsed current guidelines developed for countries with high level resources as providing excellent benchmarks for future development, and providing a logical pathway by which countries, healthcare systems and institutions can sequentially work toward better care for their populations.

References

1. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *Bmj*. 1999 Feb 20;318(7182):527-30.
2. Burgers JS, van Everdingen JJ. Beyond the evidence in clinical guidelines. *Lancet*. 2004 Jul 31-Aug 6;364(9432):392-3.
3. Agree Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Quality & Safety in Health Care*. 2003 Feb;12(1):18-23.
4. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. *JAMA*. 1999 May 26;281(20):1900-5.
5. Vlayen J, Aertgeerts B, Hannes K, Sermeus W, Ramaekers D. A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit. *International Journal for Quality in Health Care*. 2005 Jun;17(3):235-42.
6. Graham ID, Calder LA, Hebert PC, Carter AO, Tetroe JM. A comparison of clinical practice guideline appraisal instruments. *Int J Technol Assess Health Care*. 2000 Autumn;16(4):1024-38.
7. [The AGREE Collaboration. Appraisal of Guidelines for Research & Evaluation \(AGREE\) Instrument](#) 2001.
8. Graham ID, Harrison MB. Evaluation and adaptation of clinical practice guidelines. *Evid Based Nurs*. 2005 Jul;8(3):68-72.
9. Burgers JS. Guideline quality and guideline content: are they related? *Clin Chem*. 2006 Jan;52(1):3-4.
10. Harpole LH, Kelley MJ, Schreiber G, Toloza EM, Kolimaga J, McCrory DC. Assessment of the scope and quality of clinical practice guidelines in lung cancer. *Chest*. 2003 Jan;123(1 Suppl):7S-20S.
11. [Scottish Intercollegiate Guidelines Network. SIGN 50: a guideline developers' handbook. Edinburgh](#) The Network; 2004.
12. Verkerk K, van Veenendaal H, Severens JL, Hendriks EJM, Burgers JS. Considered judgement in evidence-based guideline development. *International Journal for Quality in Health Care* (In Press). 2006.
13. National Institute for Health and Clinical Excellence. [The guidelines manual](#). London: National Institute for Health and Clinical Excellence.
14. Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I, et al. What's the evidence that NICE guidance has been implemented? Results from a national evaluation

- using time series analysis, audit of patients' notes, and interviews. *Bmj*. 2004 Oct 30;329(7473):999.
15. Boluyt N, Lincke CR, Offringa M. Quality of evidence-based pediatric guidelines. *Pediatrics*. 2005 May;115(5):1378-91.
 16. Lavis JN, Posada FB, Haines A, Osei E. Use of research to inform public policymaking. *Lancet*. 2004 Oct 30-Nov 5;364(9445):1615-21.
 17. Woolf SH. Evidence-based medicine and practice guidelines: an overview. *Cancer Control*. 2000 Jul-Aug;7(4):362-7.
 18. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess*. 2004 Feb;8(6):iii-iv, 1-72.
 19. Ouimet M, Landry R, Amara N, Belkhdja O. What factors induce health care decision-makers to use clinical guidelines? Evidence from provincial health ministries, regional health authorities and hospitals in Canada. *Soc Sci Med*. 2006 Feb;62(4):964-76.
 20. Clancy CM, Cronin K. Evidence-based decision making: global evidence, local decisions. *Health Aff (Millwood)*. 2005 Jan-Feb;24(1):151-62.
 21. Christiaens T, De Backer D, Burgers J, Baerheim A. Guidelines, evidence, and cultural factors. *Scand J Prim Health Care*. 2004 Sep;22(3):141-5.
 22. Global Programme on Evidence for Health Policy. [Guidelines for WHO guidelines](#) (EIP/GPE/EQC/2003.1.). 2003 [cited 9 Aug 2006]
 23. Anderson BO, Braun S, Carlson RW, Gralow JR, Lagios MD, Lehman C, et al. Overview of breast health care guidelines for countries with limited resources. *Breast J*. 2003 May-Jun;9 Suppl 2:S42-50.

Appendix 3 - Findings from the informal survey of international guideline developers

Interviews were conducted with representatives from 11 international guideline development agencies. The interviewees were asked to provide information on the use of guideline appraisal instruments, the processes and tools they use to assess guidelines for national and regional applicability.

Use of Appraisal Instruments

The [AGREE](#) instrument has had a major impact on the way major guideline developers assess the quality of guidelines produced both internally and externally. Agencies contacted reported that the principles of the AGREE instrument have been incorporated into training handbooks and technical manuals ([NICE](#), [NZGG](#), [SIGN](#), CCO, CBO and AEZQ) and they should therefore produce guidelines that meet the AGREE criteria. NICE, SIGN and CCO do not conduct a separate AGREE type assessment of their own guidelines.

Some agencies also assess “seed” guidelines developed outside of their own organisation and sometimes outside their country or province, as the basis for their adapted local guidelines. NICE, NZGG, SIGN, CCO and CBO appraise the quality of these guidelines using the AGREE instrument. The draft ADAPTE Handbook for Guideline Adaptation⁽¹⁾ recommends that “seed” guidelines be appraised using a truncated version of the AGREE instrument - the rigour of development domain, to see if a guideline meets methodological quality standards.

GAC does not have a guideline development role, but uses the AGREE instrument to assess the applicability of a guideline for use in the family physician setting.

[COGS](#) is a checklist developed in the US. It closely resembles AGREE although it was intended to be more closely linked with usability of a guideline. COGS appears to be more widely used in the USA. The [American College of Chest Physicians](#) and [KDOQI \(Kidney Disease Outcomes Quality Initiative\)](#) use AGREE.

Both appraisal instruments (AGREE and COGS) ask questions about applicability of the guidelines (e.g., identifying organisational barriers to care, considering the costs of applying the recommendations, whether the guideline contains criteria for audit and monitoring). However, it is clear that an appraisal using one of these instruments is only one dimension of the processes used for assessing whether a guideline (even one developed internally or locally) is going to be relevant to the context of care.

Some developers in the US report that the AGREE instrument is difficult to apply and that it doesn't lend itself to composite scores and doesn't measure some of the actionable aspects of guidelines. Agencies in the US now tend to place a greater emphasis on implementation and audit and monitoring criteria and are working to find ways to assess recommendations according to their impact on health outcomes.

Both AGREE and COGS appear to be easy to use and do not require significant training. Dave Davis reported the process for training family physicians in Ontario in the use of AGREE. But it also became apparent from the discussions that the AGREE appraisal was only one dimension the family physicians were asked to comment on. Ontario GPs are encouraged to provide qualitative advice on whether they think the guideline is workable in their local context. These additional comments can have a major impact on whether the guideline goes ahead or is modified. Most other agencies either ask people within their office

or who are associated with their organisation to conduct the AGREE appraisal for externally produced, completed guidelines, or see AGREE as only one of the ways of extracting comment and feedback on draft publications.

In Australia, whether guidelines are developed by the National Health and Medical Research Council (NHMRC) or by other organisations who seek NHMRC approval of the final guideline, the process of development is guided by a multi-disciplinary committee. Whilst this includes clinicians who are likely to be involved in implementing the guideline in their own practice, their involvement is from a 'development' point of view rather than assessing the final guideline. The NHMRC has introduced a quality assurance system of independent review of all guidelines prior to approval. This is a process, not a content review. The NHMRC does not use either the AGREE or COGS instruments. The NHMRC has also produced a set of [toolkits](#) that give detailed instruction on the methodology to be followed.

In 2005, a new guideline implementability tool was launched, [GLIA](#). GLIA is now in use by the American Academy of Pediatrics and is currently being trialled in the UK by NICE. GLIA is primarily focused on assessing the implementability of individual recommendations in a guideline. It was designed as a tool to be used pre-publication so that if a recommendation fails it can be modified. The tool also identifies a number of global questions about the internal consistency and workability of the guideline publication/recommendations and its credibility. Some of the individual questions also challenge the reviewer considering the usefulness of the recommendations to consider issues such as cost and service availability etc. At this stage no further work has been done (or is planned) to test the inter-rater reliability of the tool and predictive validity of GLIA. Some individuals within the implementation directorate and guidelines team at NICE are starting to explore how the GLIA tool might be used but no formal evaluation is planned as yet.

[DISCERN](#) is used as a way of assessing the usefulness of consumer resources produced in support of a guideline. NZGG, NICE, AEZQ and CBO use this tool. However, because of the narrowness of its application it does not seem to be a relevant instrument for use as part of any planned NICS audit of Australian guidelines.

Consideration of National and Regional Contexts

Interviews disclosed that all agencies review the guidelines they develop, adapt or endorse to assess whether they will be applicable within their own particular contexts (ie national/ regional/ provincial). In most cases this is done external to appraisal using a validated guideline appraisal tool such as AGREE or COGS.

AEZQ has modified the AGREE instrument and uses the [DELBI](#). This instrument includes a domain on the applicability of the guideline to the German Healthcare System. It asks:

1. Are there recommendations for preventive, diagnostic, therapeutic and rehabilitative interventions for different areas of care?.
2. Is there information as to which interventions seem to be unsuitable, redundant or outdated?
3. Is the clinical information of the guideline organised to ensure that the process of clinical decision making is systematically presented and easily understandable?
4. Is a strategy/concept for the easy accessibility and dissemination of the guideline presented?
5. Is a concept for implementing the guideline described?
6. Is the guideline supplemented by a description of the methods used to develop the guideline?

Assessments are conducted by AEZQ staff. This instrument has been designed solely for use in Germany and has not been formally validated as yet.

NZGG asks reviewers from consumer, provider, policy and funding agencies to assess its draft guidelines. In addition to an AGREE appraisal, reviewers are asked to comment on the clinical workability of the guideline and in particular respond to the following questions:

1. Does the guideline present clear information about the management options available in New Zealand and the likely consequences of each option?
2. What changes will be necessary to implement the guideline recommendations in NZ?
3. Is the research base behind the recommendations clear and convincing? Is there any relevant evidence you know of that has not been sourced in this draft?
4. Is any additional information needed to make these guidelines useful in a NZ setting?
5. Do you have any comment on the tone, format, readability and applicability of the content for the audience that would be useful.

At Cancer Care Ontario, Canada, a new tool (PFQ) has been introduced as a way of assessing health practitioner views on a draft guideline. CCO decided that it was most important to have wide user evaluation of the draft guideline to see if health care workers would use it and to find out perceptions of the advice. The PFQ approach is seen as at variance with what an instrument like AGREE is trying to assess - that is, whether there is an absolute truth about the quality of a guideline. The PFQ is seeking advice on the areas where there is misunderstanding or lack of clarity so that the draft guideline can be improved.

While some PFQ questions relate to the methodology and the level of comfort clinicians would have using the guidelines, they also ask reviewers to complete a checklist with a 5 point Likert scale, for example:

- The draft recommendations are too rigid to apply to individual patients.
- To apply the draft recommendations will require reorganization of services and care in my practice setting.
- To apply the draft recommendations will be technically challenging.
- The draft recommendations are too expensive to apply.
- When applied, the draft recommendations will result in better use of resources than current usual practice.

CBO produces its guidelines in two parts. The first is the scientific assessment of the evidence, and the second part is a considered judgement process checklist that covers:

- clinical impact;
- safety;
- patient perspective;
- professional perspective;
- availability of services and knowledge and skills;
- organisation of care;
- costs and cost-effectiveness;
- legal aspects;
- ethical considerations; and
- industrial interests.

These "Other Considerations" were identified in response to criticisms from guideline users who said that previously the guidelines were not practical. Now these other considerations must be taken into account before a recommendation is made by the guideline development team.

These factors have been described in a paper soon to be published in the International Journal for Quality in Healthcare.⁽²⁾ The issues of applicability and implementability include an international collaboration currently drafting a manual for guideline adaptation (That is, using seed guidelines developed overseas or by a different jurisdiction and modifying those guidelines for a new country or setting). The ADAPTE group as the collaboration is known, recommends an initial assessment of possible guidelines for adaptation using the rigour dimension of the AGREE instrument. They subsequently assess each of the individual guideline recommendations for applicability and acceptability by asking the following questions:

- Does the intervention meet patient views and preferences in the context of use?
- Is the intervention/equipment available in the context of use?
- Is the necessary expertise (knowledge and skills) available in the context of use?
- Are there any constraints, organisational barriers, legislation, policies, and resources, in the healthcare setting of use that would impede the implementation of the recommendation?
- Is the recommendation compatible with the culture and values in the setting where it is to be used?
- Is the benefit from this recommendation worth implementing?⁽¹⁾

Getting Support from Policy Makers/ Funders

For all of the guideline development agencies, ensuring that the guideline was relevant to their local care setting was crucial. In addition, many guideline developers contacted receive funding for their work from government agencies. These agencies frequently require that published guidelines conform with government policies and standards. In some cases additional checklists have been developed or are used to assess whether the guidelines meet these criteria.

Costs

As well as including cost effectiveness analyses in every guideline it produces NICE is required to undertake an [economic evaluation/cost-impact](#) assessment for each technology appraisal and clinical guideline. The costing process has two main outputs:

- A national cost impact report and
- A local cost template that can be used by healthcare planners to determine the local cost of implementing the guidelines.

SIGN also has a [checklist](#) that has to be completed before the guideline is published. In its [handbook](#)⁽³⁾ it states that:

If no economic studies have been included in the evidence base, consideration should be given as to whether any recommendation based on the evidence will have resource implications for the NHS in Scotland. The questions that should be considered include:

Does the proposed course of action involve the use of resources currently available to all parts of the NHS?

- If not, are the resources likely to be available to any part of the NHS (e.g. in cities but not rural areas)?
- If the required resources are unlikely to be widely available, is there evidence to support an alternative course of action?
- Is implementation of the recommendation likely to require new resources, or the reallocation of existing resources?
- If so, how significant is the demand for additional resources likely to be? Would a full-scale economic analysis of the options be justifiable?

(Note that "resources" covers a wide range of items other than financial resources - trained staff, clinic time, specialised diagnostic or treatment facilities, etc)

NZGG is required to identify cost and service issues that would need to be considered for the full implementation of the guideline recommendations but no formal checklist has been developed. Either prior to the guideline work starting, or part way through the development process, NZGG holds discussions with the main agencies involved in the purchase or approval of pharmaceuticals, and other health officials who will decide if a formal cost impact report is required.

The NHMRC in Australia recommends that the economic feasibility of a guideline's recommendations should be assessed in accordance with their [toolkit on how to compare the costs and benefits](#) ⁽⁴⁾

Appropriateness

SIGN, NZGG, AEZQ and NICE are also required to ensure that patient concerns and wishes in relation to the treatment that is offered, are canvassed as part of the development process. This may also include consideration of cultural appropriateness. SIGN assesses the relevance of the guidelines to its local setting by ensuring that the guideline is widely reviewed by clinical and consumer groups (expand on this), and that costs arising from implementing the guideline are identified. A checklist has been developed as part of the guideline handbook. In addition all SIGN publications must be consistent with the [NHS Scotland's Equity and Diversity toolkit](#) ⁽⁵⁾ to ensure that guidelines do not disadvantage people from the following equality target groups, wherever appropriate:

- black and minority ethnic communities (including gypsy/travellers and refugees & asylum seekers);
- women and men;
- religious/faith groups;
- disabled people;
- older people;
- children and young people; and
- the lesbian, gay, bisexual and transgender community.

Guidelines are also required to recognise the impact of:

- poverty;
- mental health;
- homelessness;
- involvement in the criminal justice system;
- marital status; and
- language or social origins.

In 2005 NICE also issued similar principles relating to [Social Value Judgements](#) ⁽⁶⁾ that take account of ethical principles, preferences, culture and aspirations that underpin the nature and extent of care provided.

In New Zealand, ensuring that the guideline actively takes account of cultural considerations, as well as other special needs such as geographic isolation, and socio-economic deprivation is critical for policymakers, consumers and practitioners. NZGG has internal processes to review these and stakeholder groups and opinion leaders are consulted to ensure these elements are covered prior to publication and are a requirement of funding contracts for the guidelines being developed. At this stage a checklist has not been developed.

In Canada and Germany these factors do not appear to be explicitly considered.

In Australia, the NHMRC has recently published a paper on [Cultural Competency for Health](#) ⁽⁷⁾ and it is anticipated that cultural issues will become one of the critical issues for guideline developers in Australia in the coming years. The NHMRC produced a [toolkit on using socio-economic evidence](#) in 2002 and also advocates that guideline developers use an [equity-focused health impact assessment](#) (EFHIA)⁽⁸⁾

Performance indicators and audit criteria

All guideline developers reported that this is an area where there is great deal of interest and they see description of indicators and as an integral requirement of any guideline. This is seen as a crucial component in the US and AHRQ has developed a [quality measures/indicators clearing house](#).

CBO is in the process of developing an Appraisal of Indicators through Research and Evaluation (AIRE) instrument to assess whether the indicators included in their guidelines are both evidence-based and appropriate. An international trial of the instrument is planned for later in 2006 and Jako Burgers has agreed to notify NICS of the trial to see if there is interest in trialing the tool in Australia.

The NHMRC toolkit - [How to put evidence into practice: implementation and dissemination strategies](#),⁽⁹⁾ advocates that measurable outcomes should be included within each guideline.

Since 1st April 2006 NICE has commissioned the [Clinical Accountability, Service Planning and Evaluation](#) (CASPE) Research Unit and [Health Quality Service](#) (HQS) to develop audit criteria for all its guidelines as part of its implementation strategy.

Getting buy-in

Endorsement and support

One of the critical success factors identified by all the people interviewed was making sure that the stakeholder groups (including health care providers and consumers) had been involved in the drafting of the guideline. There is a considerable amount of literature on the science of implementing evidence and guidelines in support of this approach.

Many agencies invite stakeholders to actively review drafts of the guidelines either by posting the drafts on the website, or sending copies out to a diverse group of stakeholders including consumers, providers and funders/ policy makers, for comment. In some cases these stakeholders are asked to endorse the guidelines (eg GAC, NHMRC and AAP). Other agencies such as AEZQ and NZGG, invite stakeholder groups to attach their logos to the final guidelines. Many agencies (AEZQ, NICE, SIGN, and CBO) convene public meetings or have internet forums to seek feedback on draft guideline.

NICE reports that they have a newly established [implementation directorate](#). It convenes a workshop to consider the implementation issues arising from a draft guideline. At that workshop participants from key national organisations are asked to comment on barriers and facilitators that will affect the implementation of key recommendations. The directorate produces implementation advice and awareness raising slide-sets.

The American Academy of Pediatrics has a two stage process when guidelines are submitted for endorsement. They are first assessed using COGS and GLIA, and then they are reviewed by a policy committee to ensure the recommendations are consistent with the views of the Academy.

In many countries (Germany, Scotland, England, Canada, Netherlands, New Zealand and Australia), consumer groups are also invited to endorse and review the guideline to ensure that the guideline meets the communities of people who will benefit (or not) from the guideline.

Understandability and usability

Most agencies report that they spend more time refining the text of the guideline to ensure it is clear, unambiguous and actionable. They feel that this activity is of growing importance and requires more detailed attention and they often invite end users of the guideline to assist them with this clarification.

Some of this consultation is achieved through the administration of questionnaires and the use of tools such as GLIA, and others through a consultative process, and the development and refinement of summaries and consumer resources. This needs to be done at both the national and local level.

Some individuals within the implementation directorate and guidelines team at NICE are starting to explore how the GLIA tool might be used but no formal evaluation is planned as yet.

A number of agencies also invite an independent person/s to review the guidelines for “readability” before they are published.

Local application

All of the agencies reported that they expected that there would be local adaptation of the national guidelines and emphasise the need to ensure that the evidence in the national guideline is clearly understood so that the translation for local conditions does not alter the intent of the original recommendations.

Computerising recommendations

Computerisation of recommendations has been a big issue in the AAP and it appears to be high on the US agenda. The NGC has recently had a meeting to consider whether the Clearing House would begin quality assessing the guidelines it houses. The proposed approach is to assess the guidelines for usability and in particular - whether the recommendations can be easily turned into computerised format so that they can be integrated into electronic health record systems.

A number of other countries are developing a growing awareness of the demand for translating their guidelines into electronic formats (particularly France, the Netherlands and New Zealand).

NICE has recently started the Electronic Guidance Access Project (EGAP) as part of its implementation strategy. The project will comprise several strands to enhance the efficient & effective dissemination of the knowledge contained in NICE guidelines.

Summary Chart of the Interview Findings

	Appraisal of draft guideline prior to publication	Appraisal of published guidelines produced outside jurisdiction for adaptation or endorsement	Appraisal of supporting documentation, e.g., consumer resources
AHRQ	N/A but report that some US agencies are using both COGS and AGREE	Considering introducing an assessment of the usability and implementability of guidelines on the US Guidelines Clearing House – especially the computability of recommendations (meeting planned for Oct to consider this)	N/A
AEZQ	Reviewed using DELBI instrument (based on AGREE) plus consultation process with stakeholders and web-based review process	DELBI	DISCERN used for consumer resources
AAP	COGS statement used plus GLIA review of major recommendations plus review by policy committee	COGS statement used plus GLIA review of major recommendations plus review by policy committee	N/A
HAS	Use AGREE plus assess against criteria for improving health outcomes and access to services. Also test with 40-60 practitioners to see if recommendations feasible	Use AGREE	N/A
CBO	AGREE plus a considered judgement checklist that covers off many of the policy and practical applications. Comprehensive consultation with stakeholders also central component of guideline review - including government, insurance companies, professional groups and consumer groups.	Use AGREE - but external guidelines used to “seed” a local guideline team process and development of local recommendations	DISCERN used
CCO	Reviewed using PFQ tool (inspired by AGREE) and wide consultation with stakeholder groups	AGREE - but external guidelines used to “seed” a local guideline team process and development of local recommendations	N/A
GAC	N/A	AGREE plus qualitative comments on whether the guideline is applicable for family physician in Ontario	Consumer resources not produced by GAC does produce guideline summaries for physicians

	Appraisal of draft guideline prior to publication	Appraisal of published guidelines produced outside jurisdiction for adaptation or endorsement	Appraisal of supporting documentation e.g, consumer resources
NHMRC	NHMRC checklists and compliance with NHMRC toolkits	NHMRC checklists and compliance with NHMRC toolkits	NHMRC checklists and compliance with NHMRC toolkits
NICE	AGREE criteria included in technical manual for guideline producers, social value judgements assessed. Two independent reviews. One by the Guidelines Review Panels, the other an external peer review system by the Health Technology Assessment Programme. Cost impact assessment undertaken, review panel of clinical advisers assess guideline for usability.	AGREE for screening other guidelines- Some information may be used from other guidelines but up until now have not adapted or endorsed other guidelines	
NZGG	AGREE criteria included in guidelines handbook for internal guideline producers. Reviewers complete an AGREE appraisal plus a series of questions about the applicability of g/l for NZ context.	AGREE - but external guidelines used to “seed” a local guideline team process and the development of local recommendations	DISCERN used for consumer resources
SIGN	AGREE criteria included in guidelines handbook SIGN 50 for internal guideline producers. SIGN 50 also has checklist to examine the economic implications of the guideline. External reviewers assess usability. NHS Equality and diversity toolkit checklist used.	AGREE - but external guidelines used to “seed” a local guideline team process and the development of local recommendations	No tool used

References

1. ADAPTE group. Manual for guideline adaptation [draft]. 2006.
2. Verkerk K, van Veenendaal H, Severens JL, Hendriks EJM, Burgers JS. Considered judgement in evidence-based guideline development. *International Journal for Quality in Health Care* (In Press). 2006.
3. [Scottish Intercollegiate Guidelines Network. SIGN 50: a guideline developers' handbook. Edinburgh: The Network; 2004.](#)
4. [National Health and Medical Research Council \(Australia\). How to compare the costs and benefits : evaluation of the economic evidence. Canberra, ACT: National Health and Medical Research Council; 2001](#)
5. [Scottish Executive. Equality and diversity impact assessment toolkit. 2005 \[cited 9 Aug 2006\]](#)
6. [National Institute for Health and Clinical Excellence. Social value judgements: principles for the development of NICE guidance. 2005 \[cited 9 Aug 2006\];](#)
7. [NHMRC. Cultural competency in health: A guide for policy. partnerships and participation. Canberra: National Health and Medical Research Council; 2005.](#)
8. [Mahoney M, Simpson S, Harris E, Aldrich R, Stewart Williams J. Equity focused health impact assessment framework: the Australasian Collaboration for Health Equity Impact Assessment \(ACHEIA\); 2004. <http://chetre.med.unsw.edu.au/files/Harris-Roxas B \(2004\) Equity Focused HIA.pdf>](#)
9. [NHMRC. How to put the evidence into practice: implementation and dissemination strategies. Canberra: National Health and Medical Research Council; 2000.](#)
10. [The AGREE Collaboration. Appraisal of Guidelines for Research & Evaluation \(AGREE\) Instrument.; 2001.](#)
11. [Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. *Ann Intern Med.* 2003 Sep 16;139\(6\):493-8.](#)
12. [Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, et al. The GuideLine Implementability Appraisal \(GLIA\): development of an instrument to identify obstacles to guideline implementation. *BMC Med Inform Decis Mak.* 2005;5:23.](#)
13. [Charnock D. The DISCERN handbook: quality criteria for consumer health information on treatment choices. Oxford: Radcliffe Medical Press Ltd; 1998](#)
14. [Beyer M, Geraedts M, Gerlach FM, Guelich M, Jaeckel WH, Kopp I, et al. DELBI - German instrument for methodological guideline appraisal. 2005/2006 \[cited 9 Aug 2006\];](#)

Appendix 4- Summary of current guideline appraisal tools and instruments

This Appendix contains an outline of the three most highly regarded/ used international guideline appraisal tools (AGREE, COGS and GLIA).

To keep this document small and easy to use on the web, appraisal instruments and tools referred to in other parts of this report have been hyperlinked in the main body of the text, so that they can be accessed online.

AGREE

The AGREE ([Appraisal of Guidelines for Research and Evaluation Instrument](#)),⁽¹⁾ was developed by an international collaboration to define the basic quality criteria of a guideline. The Appraisal instrument uses a 5 point Likert scale to assess a guideline, and the developers advise that up to 4 people should be involved in appraising a single guideline. There is no cut off point that indicates a pass/fail, and many guideline developers use the AGREE instrument at the draft stage to assess whether needs modification before it is published. The AGREE instrument covers six domains.

1. Scope and purpose - Guidelines should contain a specific statement about the overall objective(s), clinical questions, and describe the target population.
2. Stakeholder involvement - Guidelines should provide information about the composition, discipline, and relevant expertise of the guideline development group and involve patients in their development. They also clearly define the target users and have been piloted prior to publication.
3. Rigour of development - Guidelines are expected to provide detailed information on the search strategy, the inclusion and exclusion criteria for selecting the evidence, and the methods used to formulate the recommendations. The recommendations are explicitly linked to the supporting evidence and there is a discussion of the health benefits, side effects, and risks. They have been externally reviewed before publication and provide detailed information about the procedure for updating the guideline.
4. Clarity and presentation - Guidelines should contain specific recommendations on appropriate patient care and consider different possible options. The key recommendations are easily found. A summary document and patients' leaflets are provided.
5. Applicability - Guidelines should discuss the organisational changes and cost implications of applying the recommendations and present review criteria for monitoring the use of the guidelines.
6. Editorial independence - Guidelines should include an explicit statement that the views or interests of the funding body have not influenced the final recommendations. Members of the guideline group have declared possible conflicts of interest.

The AGREE instrument has been validated and tested in 13 countries and hundreds of guidelines have been appraised using this instrument⁽⁶⁾ and further validation is underway as part of the AGREE II project.

Guideline developers interviewed report that they use the Rigour of Development domain as a way of quality assuring guidelines that are going to be adapted and this approach has now been incorporated into ADAPTE's draft handbook for Adapting Guidelines.⁽⁷⁾

COGS

[COGS](#) is a guideline appraisal checklist produced by a group of US guideline developers and implementers at a Conference on Guideline Standardization (COGS) in 2002.⁽²⁾ It was developed to promote the systematic reporting of precise details that are critical for understanding a guideline's development its recommendation statements and potential issues in its application. The conference defined 18 items in its checklist

- overview material;
- focus;
- goal;
- users/setting;
- target population;
- developer;
- funding source/
- evidence;
- recommendation grading criteria;
- method for synthesizing evidence;
- pre-release review;
- update plan;
- definitions;
- recommendations and rationale;
- potential benefits and
- patient preferences;
- algorithm; and
- implementation.

This checklist is used by reviewers to determine whether each of these topics is adequately covered in a guideline. It was tested by representatives from 22 US agencies active in guideline development. The checklist is currently used by a number of US guideline developers and guideline endorsement agencies. The developers note that this checklist should not be used alone to judge guideline quality or adequacy.

GLIA

In 2005, a new guideline implementability tool was launched, GuideLine Implementability Appraisal ([GLIA](#)). GLIA is primarily focused on assessing the implementability of individual recommendations in a guideline. It was designed as a tool to be used pre-publication so that if a recommendation fails it can be modified. The tool also identifies a number of global questions about the internal consistency and workability of the guideline publication/ recommendations and its credibility. Some of the individual questions also challenge the reviewer considering the usefulness of the recommendations to consider issues such as cost and service availability etc. At this stage no further work has been done (or is planned) to test the inter-rater reliability of the tool and predictive validity of GLIA. It is possible that the results from the NICE trial could assist with future enhancements of the tool (see Appendix 3).

The GLIA tool is primarily focused on assessing the implementability of a guideline. The developers define *implementability* as a set of characteristics that predict the relative ease of implementation of guideline recommendations.

Implementability is an abstract construct, relating to a number of factors, some of which are intrinsic to the guideline itself-and therefore are under the control of the developers (such as

the clarity and lack of ambiguity or incompleteness of the recommendations ie whether the people know what to do, how and when).

The tool also highlights extrinsic factors that are largely site-specific but relate to the recommendations on the effects of the process of care and relating to novelty or innovation of a guideline statement (such as the availability of equipment and practitioner skills, or that the recommendations might not be consistent with patient or consumer expectations). Consideration of these issues is required if the guideline is to be used locally so that developers can anticipate barriers and propose potential strategies for implementation success.

Measures of successful implementability focus on the ease and accuracy of translating the guidelines into systems that influence care. While the tool considers some general global considerations of the guideline construction (7 questions) and, the majority of the tool (an additional 24 questions) are directed to the usability of individual recommendations.

Intrinsic Factors identified in GLIA

- **Executability** (exactly what to do under the circumstances defined)
- **Decidability** (precisely under what circumstances to do something)
- **Presentation and formatting** (the degree to which the recommendation is easily recognizable and succinct)
- **Measurable outcomes** (the degree to which the guideline identifies markers or endpoints to track the effects of implementation of this recommendation)
- **Apparent validity** (the degree to which the recommendation reflects the intent of the developer and the strength of evidence)

Extrinsic Factors identified in GLIA

- **Effect on Processes of Care** (the degree to which the recommendation impacts on the usual workflow of a care setting)
- **Novelty/innovation** (the degree to which the recommendation proposes behaviours considered unconventional by clinicians or patients)
- **Flexibility** (the degree to which a recommendation permits interpretation and allows for alternatives in its execution)
- **Computability** (the ease with which a recommendation can be operationalized in an electronic information system) is only applicable when an electronic implementation is planned.

The aim of the instrument is not to establish a cut-off point or pass/ fail for a guideline, but to identify where modification or improvement of a guideline is required. It is recommended by the GLIA developers that two or more people review a guideline at any one time. An electronic version of the instrument eGLIA⁽⁸⁾ has been developed to allow reviewers from around the US to simultaneously rate a draft guideline.

References

1. [The AGREE Collaboration. Appraisal of Guidelines for Research & Evaluation \(AGREE\) Instrument.; 2001.](#)
2. [Shiffman RN, Shekelle P, Overhage JM, Slutsky J, Grimshaw J, Deshpande AM. Standardized reporting of clinical practice guidelines: a proposal from the Conference on Guideline Standardization. Ann Intern Med. 2003 Sep 16;139\(6\):493-8.](#)

3. [Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, et al. The GuideLine Implementability Appraisal \(GLIA\): development of an instrument to identify obstacles to guideline implementation. BMC Med Inform Decis Mak. 2005;5:23.](#)
4. [Charnock D. The DISCERN handbook: quality criteria for consumer health information on treatment choices. Oxford: Radcliffe Medical Press Ltd; 1998.](#)
5. [Beyer M, Geraedts M, Gerlach FM, Guelich M, Jaeckel WH, Kopp I, et al. DELBI - German instrument for methodological guideline appraisal. 2005/2006 \[cited 9 Aug 2006\];](#)
6. Burgers JS, Fervers B, Haugh M, Brouwers M, Browman G, Philip T, et al. International assessment of the quality of clinical practice guidelines in oncology using the Appraisal of Guidelines and Research and Evaluation Instrument. J Clin Oncol. 2004 May 15;22(10):2000-7.
7. ADAPTE group. Manual for guideline adaptation [draft]. 2006.
8. Shiffman R, Codish S. The PPI Toolkit and eGLIA