

A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit

JOAN VLAYEN^{1,2}, BERT AERTGEERTS^{1,3}, KARIN HANNES¹, WALTER SERMEUS² AND DIRK RAMAEKERS^{1,4}

¹Belgian Centre for Evidence-Based Medicine, Leuven, Belgium ²Catholic University Leuven, Centre for Health Services and Nursing Research, Leuven, Belgium ³Catholic University Leuven, Academic Centre for General Practice, Leuven, Belgium, and ⁴Belgian Federal Health Care Knowledge Centre, Brussels, Belgium

Abstract

Objective. To identify a critical appraisal tool for clinical practice guidelines that could serve as a basis for the development of an appraisal tool for clinical pathways.

Design. Systematic review of the literature and personal contacts. Databases searched were: Medline, Embase, and Cinahl. Search terms were: practice guidelines, appraisal, and evaluation. The items of the identified appraisal tools were examined and thematically grouped into 10 guideline dimensions. Content analysis and scoring of these domains by the appraisal tools was evaluated.

Results. Twenty-four different appraisal tools of practice guidelines were identified. None scored the evidence base of the clinical content of guidelines. Four tools scored all the guideline dimensions. The Cluzeau instrument is the only one of these four that has been validated. Of the three instruments based on the Cluzeau instrument, the AGREE instrument is the only validated instrument that uses a numerical scale.

Conclusions. Being a simplified version of the Cluzeau instrument, the AGREE instrument has the most potential to serve as a basis for the development of an appraisal tool for clinical pathways. However, important limitations will have to be dealt with when developing such a tool.

Keywords: appraisal tool, clinical pathways, content analysis, practice guidelines

In 1990, the Institute of Medicine (IOM) defined clinical practice guidelines as systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances [1]. Great variability exists in the quality of clinical practice guidelines [2,3], and numerous appraisal instruments were therefore developed in an attempt to discriminate high-quality from lower-quality guidelines. A recent review of the literature identified 13 appraisal instruments of practice guidelines, but concluded that evidence at that time was insufficient to support the exclusive use of just one appraisal instrument [4].

Selection of a decent guideline is one thing, but the selected guideline still has to be implemented. One way to implement clinical practice guidelines into daily practice is through clinical pathways. These are sequences of standardized multidisciplinary processes or critical interventions that must occur for a specific population towards the desired outcomes within a defined time period [5]. Originally, nursing care, organization, and cost-effectiveness were the most important aspects to be

covered by clinical pathways. However, they also provide a means of improving systematic collection and abstraction of clinical data for audit and of promoting change in practice [6]. Increasingly, questions on the ability of clinical pathways to improve clinical quality, patient satisfaction, and the satisfaction of care providers are being researched.

Furthermore, it is unclear whether a relationship exists between the methodological and content quality of clinical pathways on the one hand and clinical quality, defined by the judicious and explicit use of the evidence from clinical trials, on the other. The instauration of several diagnostic and therapeutic interventions in a local clinical pathway based purely on the opinion of local specialists and without systematically taking the available up-to-date evidence for these interventions into consideration, can threaten the appropriateness and quality of care.

Clinical pathways show important similarities to clinical practice guidelines. They are both intended to provide appropriate and effective health care for specific clinical circumstances and to reduce variation in practice [1,6]. However,

Address reprint requests to Joan Vlayen, Centre for Health Services and Nursing Research, Kapucijnenvoer 35, B-3000 Leuven, Belgium. E-mail: joan.vlayen@med.kuleuven.ac.be

clinical pathways commonly integrate key aspects from several clinical practice guidelines, because they outline the care to be given for the patient's entire clinical path rather than for one specific clinical situation [7]. Clinical practice guidelines are usually developed by government agencies, institutions, or expert panels, whereas clinical pathways are more local initiatives.

Unlike clinical practice guidelines, no (validated) instrument exists to assess the methodological quality of clinical pathways. A possible strategy to develop a critical appraisal tool for clinical pathways is to base it on an appraisal tool for clinical practice guidelines [7]. In the present article, the first phase of the development of such an appraisal tool for clinical pathways is described. In this first phase, we performed an update of the above-mentioned study of Graham *et al.* [4]. A systematic review of the literature was carried out to identify and compare existing critical appraisal tools for clinical practice guidelines. Possible scientific validation and international dissemination of the appraisal instruments were of particular concern in our review. In a later phase, one or more instruments will be selected and applied to clinical pathways, and serve as a basis for the development of an appraisal tool for clinical pathways.

Methods

Search strategy

A literature search of the English and non-English literature indexed in the Ovid–Medline database (1966–October 2003), Embase database (1990–October 2003), and Cinahl database (1982–October 2003) was conducted, using the following MeSH and text terms in combination: practice guidelines, appraisal, evaluation. Methodology filters were not used, in order to conduct a sensitive literature search and because the usefulness of such filters is unclear for this particular subject. A manual search of the references of relevant articles was conducted. We also contacted the developers of the instruments identified to determine whether they were aware of any other instruments to include in our review.

All articles that described the evaluation of clinical practice guidelines or the development of a guideline appraisal tool were included. Articles describing a practice guideline or the development process of a practice guideline were excluded. No restriction was placed on abstracts, conference proceedings, or language. One investigator (J.V.) assessed the selected papers and retrieved all the different tools that appraised the quality of clinical practice guidelines. Tools based completely on, or copied from, an existing tool were excluded.

Tools evaluation

The tools were compared for their general characteristics (source, items, scoring system). We also analysed whether items of specific importance in the development of clinical practice guidelines and clinical pathways were evaluated. One investigator (J.V.) therefore examined the questions/statements from all the instruments and thematically grouped them into separate guideline dimensions. Two other investigators (D.R., B.A.) independently performed a content analysis of the different instruments to

assess whether or not they covered one or more items of each dimension. The scoring of the evidence base of the clinical content was also assessed. In cases of disagreement, the instruments were analysed and discussed in a small group (J.V., D.R., B.A.).

Results

The Embase search yielded 11 523 articles of which 42 were selected. Of the Medline search another 41 of 2863 retrieved articles were selected. The Cinahl database contained 1187 articles, four being selected. A manual search of the bibliographies of the articles retrieved yielded another 11 relevant articles.

In the 98 articles retrieved, we identified 24 possible critical appraisal tools of guidelines. One tool was excluded from the review because it concerned an automated version [8] of another tool [9]. An additional instrument, not found through the systematic review of the literature or our personal contacts, was handled by one of the reviewers of this article [10].

General characteristics of the appraisal tools

Table 1 provides an overview of the characteristics of the 24 tools. The first appraisal tool was published in 1992 by Lohr and Field [11]. Before 1995, tools were exclusively published in North America [11–16]. Since 1995, appraisal tools have gained interest all over the world. Twenty-two tools were developed in eight different countries: six tools in the USA [11,12,15–18], five in Canada [13,14,19–21], four in the UK [9,22–24], two in Australia [25,26] and Italy [27,28], one in France [29], Germany [30], and Spain [31]. Two tools have been developed internationally [10,32].

Eleven instruments [9,10,13,18,21,22,24,26,27,29,32] were based on the instrument developed by the IOM [11], three instruments [9,13,19] referred to Hayward *et al.* [14], another three instruments [23,30,32] referred to Cluzeau *et al.* [22].

The number of questions ranged from 3 to 52. Some questions were subdivided into two or more smaller questions. Nine tools used no specified scoring system, 10 tools used a yes/no score with or without the possibility of answering 'not sure' or 'not applicable' (Table 1). Five instruments used some kind of scaling system [17,18,25,26,32]. The instrument developed by Sanders [18] and the AGREE instrument [32] use a numerical scale.

All but three instruments were published in peer-reviewed literature. Only four instruments have been subject to a validation study: the instruments developed by Shaneyfelt [9], Cluzeau [22], and Grilli [28], and the AGREE instrument [32].

Content analysis

In total, a list of 469 questions or statements was generated from the 24 instruments. The common questions/statements were grouped into 50 different items (Table 2). These 50 items were then grouped into 10 guideline dimensions based on the work of the IOM [33] and the study of Graham *et al.* [4] (Table 2): validity, reliability/reproducibility, clinical applicability, clinical flexibility, multidisciplinary process, clarity, scheduled review, dissemination, implementation, and evaluation. All 50 items could be fit into the 10 dimensions.

Table 1 Characteristics of critical appraisal tools of guidelines

Author ^a	Date	Country of origin	Published in peer-reviewed literature	Validation	Scoring system ^b	No. of items
Institute of Medicine [11]	1992	USA	Yes	Not stated	Y/N/NA	46
Hayward <i>et al.</i> [14]	1993	Canada	Yes	Not stated	None	9
Selker [12]	1993	USA	Yes	Not stated	None	7
Hayward <i>et al.</i> [13]	1995	Canada	Yes	Not stated	None	10
Mendelson [15]	1995	USA	Yes	Not stated	None	8
Woolf [16]	1995	USA	Yes	Not stated	None	10
SIGN [24]	1995	UK	No	Not stated	Y/N	52
Mutter-Pilson [29]	1995	France	Yes	Not stated	Y/N/NA	18
Ward and Grieco [26]	1996	Australia	Yes	No	Scale	18
Liddle <i>et al.</i> [25]	1996	Australia	No	Not stated	Scale	14
Savoie <i>et al.</i> [21]	1996	Canada	No	Not stated	Y/N	15
Calder <i>et al.</i> [19]	1997	Canada	Yes	No	Y/N	24
Shaneyfelt <i>et al.</i> [9]	1998	UK	Yes	Yes	Y/N	25
Helou and Ollenschlager [30]	1998	Germany	Yes	Not stated	Y/N/?/NA	41
Apolone and Bamfi [27]	1999	Italy	Yes	Not stated	None	6
Cluzeau <i>et al.</i> [22]	1999	UK	Yes	Yes	Y/N/?/NA	37
Grilli <i>et al.</i> [28]	2000	Italy	Yes	Yes	Y/N	3
Casi <i>et al.</i> [31]	2000	Spain	Yes	No	Y/N	21
Marshall [20]	2000	Canada	Yes	Not stated	None	9
Sanders <i>et al.</i> [18]	2000	USA	Yes	Not stated	Scale	15
Reed <i>et al.</i> [17]	2000	USA	Yes	Not stated	Scale	33
Hutchinson <i>et al.</i> [23]	2003	UK	Yes	Not stated	None	5
AGREE collaboration [32]	2003	Europe	Yes	Yes	Scale	23
Shiffman <i>et al.</i> [10]	2003	North America/UK	Yes	No	None	18

^aSIGN: Scottish Intercollegiate Guidelines Network; IMCARE: Internal Medicine Center to Advance Research and Education; APA: American Psychological Association; AGREE: Appraisal of Guidelines Research and Evaluation.

^bY: yes; N: no; NA: not applicable; ?: not sure.

The two independent reviewers of the different instruments agreed completely on four instruments [9,22,23,26]. Disagreement existed on 0–5 dimensions per instrument (mean 2.2). The dimension clinical flexibility most frequently caused disagreement (10 instruments).

All the instruments evaluate the validity of guidelines with at least one item, almost all evaluate the clinical applicability (Table 3). Approximately 75% of the instruments score the dimensions reliability/reproducibility, clinical flexibility, scheduled review, and multidisciplinary process. Fourteen tools address dimension clarity. Only a minority scores the dimensions dissemination, implementation, and evaluation. Three appraisal tools (Table 3) score all the dimensions mentioned above by using at least one question [22,24,30]. None of the instruments score the evidence base of the clinical content of guidelines.

Discussion

In addition to the 13 instruments found by Graham *et al.* [4], we identified another 11 different tools for the critical appraisal of clinical practice guidelines. Comparison of these 24 instru-

ments showed a wide variation in source, number of items, ways of scoring, and specific aspects that are scored. The questions used in the appraisal tools were grouped into 50 different items, which is slightly more than in the study of Graham *et al.* [4]. However, these 50 items could easily be fitted into the 10 guideline dimensions used by Graham *et al.*

Three appraisal tools were found to address all the guideline dimensions [22,24,30]. Of these, the Cluzeau instrument [22] is the only instrument that has been subject to a thorough validation study. It was originally based on the instrument developed by the IOM [11] and contains 37 items divided into three dimensions: rigour of development (questions 1–20), context and content (questions 21–32), and application (questions 33–37). A yes/no score is used to respond to each question.

Three additional instruments are based on the Cluzeau instrument [23,30,32]. Of these, the AGREE instrument [32] is the only instrument that has been validated. It uses a numerical scoring scale, making it easier to compare scores. It is more compact than the Cluzeau instrument, containing only 23 items divided into six domains: scope and purpose, stakeholder involvement, rigour of development, clarity and presentation, applicability, and editorial independence. Unlike

Table 2 Items of the appraisal instruments grouped into 10 guideline dimensions

Dimension	Item	Definition
Validity	Decision making: how consensus was reached	Method(s) used to reach consensus about guideline recommendation; role of values
	Decision making: how recommendations were made	Method(s) used in formulating recommendations
	Evidence collection	How the evidence was obtained
	Literature search	How the literature was searched, including search strategy
	Sources of evidence	Sources of evidence (textbooks, periodical literature)
	References cited	References for the evidence upon which the guideline was based
	Literature selection	Criteria used to in- and exclude literature from the data synthesis
	Evaluation of evidence	How the evidence was graded, which may or may not include a statement about the strength of evidence
	Synthesis of data	Method(s) by which the evidence was synthesized
	Recommendations and their evidence	Recommendations consistent with each other and the evidence used to support them
	Major recommendations	Differentiating major from other recommendations
	Links strength of evidence to recommendation	Links strength of evidence to recommendation
	Other guidelines	Existence of other guidelines relevant to guideline topic checked and compared
	Consistent with policy of guideline development organization	Consistent with policy of guideline development organization
	Alternatives	Alternative interventions to those recommended or dealt with by the guideline to deal with topic
	Health benefits	Expected health benefits of guideline mentioned
	Harms, risks	Potential harms or risks of guideline mentioned
Costs	Economic and other cost outcomes of guideline mentioned	
Outcomes stated	Outcomes expected to result from guideline stated	
Reliability/ reproducibility	Independent review	Peer review; guideline sent to experts not involved in its development for review
	Pilot/pretesting	Guideline piloted or pretested in clinical setting before its dissemination
	Documentation	Process of guideline development documented
Clinical applicability	Purpose	Goal or objective of the guideline
	Rationale	Rationale of or reason for the guideline
	Guideline topic	Topic or health problem dealt with
	Patient population	Patient population for whom the guideline is intended
	Provider population	Group of health care providers to whom the guideline is directed or who should use the guideline
Clinical flexibility	In-/outpatient	Discriminating between in- and outpatients
	Ethical aspects	Ethical aspects
Clinical flexibility	Exceptions/flexibility	Flexibility in the application of the guideline, or situations in which guidelines may not apply
	Patient preferences considered	Whether patient choices and/or views were considered

Table 2 *continued*

Dimension	Item	Definition
Clarity	Unambiguous	Guideline is clearly worded
	Presentation	Guideline presentation is user friendly
	Ease of use	Guideline can be used in a straightforward manner
	Structured abstract	Structured abstract or summary provided
	Patient information	Patient information included
Scheduled review	Scheduled review	Date guideline becomes no longer valid or is scheduled for review
	Date of issue of guideline	Date of issue of guideline
Development team	Multidisciplinary process	All relevant disciplines involved
	Composition of guideline development team	The individuals and/or disciplines, occupations, or organizations represented in the group who developed the guideline
	Conflict of interest	Consideration of any (potential) bias, (potential) conflicts of interest related to the individuals developing the guideline
	Funding and related bias	Sources of funding
	Endorsers	Endorsement of guideline by official bodies
	Guideline development organization	The organization or group who developed the guideline
Implementation	Implementation	Strategies to implement the guideline
	Feasibility	Policy and administrative implications of using the guideline
Dissemination	Dissemination	How the guideline is to be distributed to intended users
Evaluation	Evaluation	How the guideline is to be evaluated once it has been implemented
	Adherence	Adherence to the guideline by the intended users

the Cluzeau instrument, the dimension dissemination is not scored in AGREE. Because the AGREE instrument is a validated, easy-to-use, and transparent instrument, which was internationally developed and widely accepted, it can possibly serve as a basis for an instrument to evaluate the methodological quality of clinical pathways. English investigators have already reported on an appraisal tool, the Integrated Care Pathway Appraisal Tool (ICPAT) [7], which is based on the AGREE instrument, but is yet to be validated.

There are some important limitations in the use of the AGREE instrument. 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. The use of a systematic methodology in the retrieval of evidence supporting guideline development is frequently scored by appraisal tools [9,11,13,14,16,17,19–22,24–26,30–32]. However, even by using a recent and advanced instrument such as the Cluzeau instrument or the AGREE instrument, the results of the search for evidence, the correct use of inclusion and exclusion criteria, and the critical appraisal of the retrieved evidence are not validated. Therefore, a major conclusion of this review is that in order

to evaluate the quality of the clinical content and more specifically the evidence base of a clinical practice guideline, verification of the completeness and the quality of the literature search and its analysis has to be added to the process of validation by an appraisal instrument. Experience with the methodologies of evidence-based medicine such as literature search and critical appraisal is therefore essential for guideline validators to assure the quality of the appraisal process. Blind application of an appraisal instrument, even when validated and widely implemented, without particular attention to the evidence supporting the guideline, can threaten the credibility of these instruments and the current evolution in the international community to further elaborate the quality of guideline development [34].

Because of the differences between clinical pathways and clinical practice guidelines, the AGREE instrument cannot be applied to clinical pathways using the present version. Some items will have to be reformulated or removed, new items will have to be included. For example, the language used in the AGREE instrument is clearly 'guideline language' and will have to be translated into 'pathway language': for example, clinical pathways contain concrete interventions rather than recommendations, clinical pathways are implemented rather than published. At present, our team is composing a development group consisting of experts in clinical pathway development and experts in clinical practice guideline development

Table 3 Dimensions covered by the critical appraisal tools

Author	Validity	Reliability/ Reproducibility	Clinical applicability	Clinical flexibility	Clarity	Scheduled review	Development team	Implementation	Dissemination	Evaluation
Institute of Medicine [11]	X	X	X	X	X	X	X			
Hayward <i>et al.</i> [14]	X	X	X	X	X	X	X			
Selker [12]	X	X	X	X	X	X				
Hayward <i>et al.</i> [13]	X	X	X	X	X	X	X			
Mendelson [15]	X	X	X	X	X	X	X		X	
Woolf [16]	X	X	X	X	X	X	X		X	X
SIGN [24]	X	X	X	X	X	X	X			
Mutter-Pilson [29]	X	X	X	X	X	X	X			
Ward and Grieco [26]	X	X	X	X	X	X	X			
Liddle <i>et al.</i> [25]	X	X	X	X	X	X	X		X	X
Savoie <i>et al.</i> [21]	X	X	X	X	X	X	X		X	X
Calder <i>et al.</i> [19]	X	X	X	X	X	X	X			
Shaneyfelt <i>et al.</i> [9]	X	X	X	X	X	X	X		X	X
Helou and Ollenschlager [30]	X	X	X	X	X	X	X		X	X
Apolone and Bamfi [27]	X	X	X	X	X	X	X		X	X
Cluzeau <i>et al.</i> [22]	X	X	X	X	X	X	X		X	X
Grilli <i>et al.</i> [28]	X	X	X	X	X	X	X		X	X
Casi <i>et al.</i> [31]	X	X	X	X	X	X	X		X	X
Marshall [20]	X	X	X	X	X	X	X		X	X
Sanders <i>et al.</i> [18]	X	X	X	X	X	X	X		X	X
Reed <i>et al.</i> [17]	X	X	X	X	X	X	X		X	X
Hutchinson <i>et al.</i> [23]	X	X	X	X	X	X	X		X	X
AGREE collaboration [32]	X	X	X	X	X	X	X		X	X
Shiffman <i>et al.</i> [10]	X	X	X	X	X	X	X		X	X

and validation to create a version of the AGREE instrument applicable to clinical pathways. The development of this instrument will be described in a subsequent publication.

In conclusion, 24 different appraisal tools of clinical practice guidelines were identified. Of these tools, the Cluzeau instrument seems to be the most complete. Being a more compact version of the Cluzeau instrument and using a numerical scale, the AGREE instrument has the potential to serve as a basis for a critical appraisal tool for clinical pathways. However, some important limitations of the AGREE instrument will have to be dealt with when developing such a tool.

References

- Institute of Medicine Committee to Advise the Public Health Service on Clinical Practice Guidelines. *Clinical Practice Guidelines: Directions For A New Program*. Washington DC: National Academy Press, 1990.
- Saturno PJ, Medina F, Valera F, Montilla J, Escolar P, Gascon JJ. Validity and reliability of guidelines for neck pain treatment in primary health care. A nationwide empirical analysis in Spain. *Int J Qual Health Care* 2003; **15**: 487–493.
- Christiaens T, De Backer D, Burgers J, Baerheim A. Guidelines, evidence, and cultural factors. *Scand J Prim Health Care* 2004; **22**: 141–145.
- 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; **16**: 1024–1038.
- Nyberg D, Marschke P. Critical pathways: tools for continuous quality improvement. *Nurs Adm Q* 1993; **17**: 62–69.
- Campbell H, Hotchkiss R, Bradshaw N, Porteous M. Integrated care pathways. *Br Med J* 1998; **316**: 133–137.
- de Luc K, Whittle C. An integrated care pathway appraisal tool: a 'badge of quality'. *Journal of Integrated Care Pathways* 2002; **6**: 13–17.
- Agrawal A, Shiffman RN. Evaluation of guideline quality using GEM-Q. *Medinfo* 2001; **10**: 2–101.
- 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; **281**: 1900–1905.
- 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; **139**: 493–498.
- Lohr K, Field M. eds. A provisional instrument for assessing clinical practice guidelines. In *Guidelines for Clinical Practice: From Development to Use*. Washington DC: National Academy Press, 1992.
- Selker HP. Criteria for adoption in practice of medical practice guidelines. *Am J Cardiol* 1993; **71**: 339–341.
- Hayward RS, Wilson MC, Tunis SR, Bass EB, Guyatt G. Users' guides to the medical literature. VIII. How to use clinical practice guidelines. A. Are the recommendations valid? The Evidence-Based Medicine Working Group. *JAMA* 1995; **274**: 570–574.
- Hayward RS, Wilson MC, Tunis SR, Bass EB, Rubin HR, Haynes RB. More informative abstracts of articles describing clinical practice guidelines. *Ann Intern Med* 1993; **118**: 731–737.
- Mendelson EB. The development and meaning of appropriateness guidelines. *Radiol Clin North Am* 1995; **33**: 1081–1084.
- Woolf SH. Practice guidelines: what the family physician should know. *Am Fam Physician* 1995; **51**: 1455–1463.
- Reed GM, McLaughlin CJ, Newman R, American Psychological Association. American Psychological Association policy in context. The development and evaluation of guidelines for professional practice. *Am Psychol* 2002; **57**: 1041–1047.
- Sanders GD, Nease RF, Jr, Owens DK. Design and pilot evaluation of a system to develop computer-based site-specific practice guidelines from decision models. *Med Decis Making* 2000; **20**: 145–159.
- Calder L, Hébert P, Carter A, Graham I. Review of published recommendations and guidelines for the transfusion of allogeneic red blood cells and plasma. *CMAJ* 1997; **156**: S1–S8.
- Marshall JK. A critical approach to clinical practice guidelines. *Can J Gastroenterol* 2000; **14**: 505–509.
- Savoie I, Green C, Bassett K et al. *Critical Appraisal Criteria for Clinical Practice Guidelines*. Vancouver: BC Office of Technology Assessment, 1996.
- Cluzeau FA, Littlejohns P, Grimshaw JM, Feder G, Moran SE. Development and application of a generic methodology to assess the quality of clinical guidelines. *Int J Qual Health Care* 1999; **11**: 21–28.
- Hutchinson A, McIntosh A, Anderson J, Gilbert C, Field R. Developing primary care review criteria from evidence-based guidelines: coronary heart disease as a model. *Br J Gen Pract* 2003; **53**: 690–696.
- Scottish Intercollegiate Guidelines Network (SIGN). *Clinical Guidelines: Criteria for Appraisal for National Use*. Edinburgh: Royal College of Physicians, 1995.
- Liddle J, Williamson M, Irwig L. *Method for Evaluating Research Guidelines Evidence (MERGE)*. Sydney: New South Wales Department of Health, 1996.
- Ward JE, Grieco V. Why we need guidelines for guidelines: a study of the quality of clinical practice guidelines in Australia. *Med J Aust* 1996; **165**: 574–576.
- Apolone G, Bamfi F. A model for an evaluation of clinical guidelines (in Italian). *Ric Pract* 1999; **15**: 116–124.
- Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. *Lancet* 2000; **355**: 103–106.
- Mottur-Pilson C. Internists' evaluation of guidelines: the IMCARE Practice Guidelines Network. *Int J Qual Health Care* 1995; **7**: 31–37.
- Helou A, Ollenschlager G. Goals, possibilities and limits of quality evaluation of guidelines. A background report on the user manual of the 'Methodological Quality of Guidelines' check list (in German). *Z Arztl Fortbild Qualitatssich* 1998; **92**: 361–365.

31. Casi CA, Aizpuru BF, Ibanez PF, Grupo de Trabajo sobre Calidad de las Guías de Práctica Clínica en Atención Primaria de la Sociedad Vasca de Medicina de Familia. Quality analysis of clinical practice guidelines on cardiovascular risks in the Basque country and a proposal for their evaluation (in Spanish). *Aten Primaria* 2000; **26**: 585–589.
32. AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Saf Healthb Care* 2003; **12**: 18–23.
33. Institute of Medicine. *Guidelines for Clinical Practice: From Development to Use*. Washington, DC: National Academy Press, 1992.
34. Guidelines International Network: <http://www.g-i-n.net/> Accessed 2 October 2004.

Accepted for publication 5 January 2005