

Production and quality of clinical practice guidelines (CPG) in Argentina (1994-2004): a cross-sectional study.

María Eugenia Esandi^{1§}, Zulma Ortiz¹, Evelina Chapman¹, Marcelo García Dieguez^{1,2}, Mejía Raúl³, Ricardo Berzstein⁴.

¹Instituto de Investigaciones Epidemiológicas, Academia Nacional de Medicina, Ciudad Autónoma de Buenos Aires, Argentina.

²Area de Estudio de los determinantes epidemiológicos de la salud, Universidad Nacional del Sur, Bahía Blanca, Argentina.

³Programa de Medicina Interna General, Hospital de Clínicas, Universidad de Buenos Aires, Argentina.

⁴Hospital de **Pediatría "Prof. Dr. Juan P. Garrahan"**, Buenos Aires, Argentina.

[§]Corresponding author. Email address:

María Eugenia Esandi: eesandi@infovia.com.ar

Abstract

Background: In the last decades, a sustained increment of Clinical Practice Guidelines (CPG) production in the world has been accompanied by a growing concern about their quality. Many studies related to quality assessment of guidelines produced in developed Countries were published; however, evidence on this topic is scarce in Low and Middle Income Countries (LMIC). The objectives of this research were: a) to describe guideline production in Argentina at different levels of the health system (macro, meso and micro) from 1994 to 2004; and b) to assess their quality by using the AGREE instrument. **Methods:** A cross-sectional study was undertaken to describe guidelines production in Argentina between 1994 and 2004. A two step approach was used to identify Argentine guidelines by searching Internet and electronic databases (MEDLINE and LILACS). Each CPG was independently assessed by two reviewers using the AGREE instrument. Domain scores were calculated as recommended by the AGREE Collaboration. We evaluated the instrument' reliability calculating the α Cronbach. Feasibility of the instrument was analyzed by the Intraclass Correlation Coefficient (ICC). **Results:** A total amount of 431 potential CPG were identified, but only 144 were considered CPG. At the end, only 101 CPG were included for further assessment. Median standardized score for each domain were: scope = 39%; stakeholder involvement = 13%; rigor of development = 10%; clarity =42%; applicability = 6%; editorial independence = 0%. Only 22 % (22) CPG were recommended with modifications by both appraisers. ICC and α Cronbach for each domain were in all cases moderate or high (greater than 0.40), except for editorial independence. **Conclusions:** This study has sistematically employed the AGREE instrument for the critical assessment of guidelines produced in a LMIC. Even when production of guidelines in Argentina in the last eleven years showed a constant increment, quality of reporting has not improved and in some aspects, seems to be worst. Much room for improvement of the guideline development process was found at all levels of the health system.

Background

Clinical Practice Guidelines (CPG) are one of the tools most frequently used by health professionals to improve the micro level decision-making process. As defined by the Institute of Medicine (IOM), they are "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances".¹ Guidelines may offer concise instructions on which diagnostic or screening tests need to be order, how to provide medical or surgical services, how long patients should stay in hospital, or other details of clinical practice.² The ultimate purpose of developing and using guidelines is to improve the quality of care provided, particularly in areas of clinical uncertainty.

In the last years, a sustained increment in guidelines production was observed all over the world, especially in United States, Canada, Australia, New Zealand and European countries. Most of these countries have developed national programs for CPG production, dissemination and implementation in order to increase the effectiveness and quality of the health system.³

Some of these initiatives, which were originally conceived as individual efforts, have been strongly improved by international cooperation through organizations such as the Guidelines International Network (GIN)⁴ and the Appraisal of Guidelines for Research and Evaluation (AGREE) Collaboration.⁵ The need for harmonizing and systematizing guideline development and assessment was one of the most important reasons that prompted the establishment of these international organizations.

A "good quality guideline" is that one that eventually leads to improved patient outcome. However, as this type of evidence is rarely available, quality of guideline is indirectly measured by assessing in what degree guideline producers minimized potential biases that could occur in the development process and affect validity of its recommendations.⁶ Wrong recommendations affect health professionals' credibility on guidelines, and consequently, limit their adoption.²

In 1999, Shaneyfelt y col. assessed quality of CPG published in Medline between 1985 and 1997 by using an ad-hoc instrument. The majority of 279 assessed guidelines did not meet the pre-established methodological standards, being rigor of recommendations one of the most deficiently reported.⁷ Similar results were reported by Cluzeau and col.⁸, Grilli and col.⁹ and Graham and col.¹⁰ in 1999, 2000 y 2001, respectively. In 2003, the AGREE collaboration published the results of the first project aimed at developing and validating a generic instrument for guidelines assessment.⁶ This instrument has been translated to different languages, extending its use throughout the world. In the recent years, several studies showed methodology deficiencies in guideline development by using the AGREE instrument.¹¹⁻¹²⁻¹³

Almost all research about quality assessment of CPG has been undertaken in High Income Countries. Evidence of quality of guidelines produced and diffused in Low and Middle Income Countries (LMIC), and particularly in Latinoamerica, is scarce.¹⁴ In Argentina, although many different institutions are interested in CPG development, there is no evidence about quantity of guidelines produced, and moreover, quality of these documents. The purpose of this research is to describe trends in guidelines production in Argentina and to assess their quality by using the AGREE instrument.

Methods

A cross-sectional study was undertaken to describe guidelines production in Argentina between 1994 and 2004. A two step approach was used to identify Argentine guidelines by searching Internet and electronic databases (MEDLINE and LILACS). The searches were performed following the methodology reported by Marín and col.¹⁵ Search execution was performed by an expert through the strategy described in table 1. It was initially developed to be performed in MEDLINE under PubMed platform. Afterwards it was adapted to be used in regional databases (LILACS). All retrieved articles were assessed by the principal investigator in order to select those that met pCPG criteria. An Internet Searching (IS) was also performed through the identification, selection and assessment of the Websites of health institutions at the macro, meso and micro level. A systematized strategy was applied in order

to identify guidelines posted on Websites: 1. Search of sections or/and documents that fulfilled pCPG definition; 2. Search by means of Website Internal Browsers using the same key words. All those documents that fulfil pCPG criteria were registered in a database. Each of the pCPG was assessed by two of the researchers in order to select those that met CPG criteria.

The first step was aimed at identifying potential CPG (pCPG), which were defined as “any document that contains any of the following key words: *Guideline; Clinical Practice Guideline; Handling Guideline; Consensus; Recommendations; Protocol; Algorithm; Expert Opinion*”. In the second step, inclusion and exclusion criteria were applied on pCPG and a subset of CPG was obtained. Documents were considered as CPG if: 1) They included explicit recommendations targeted to health professional decision-making, being this related to: screening and primary prevention, diagnosis, treatment and secondary prevention and/or rehabilitation; 2) They contained bibliographic references and in the case of consensus, participants or responsible institutions were described; 3) They were produced and diffused in the period of study (January 1994- December 2004) and were publicly accessible. Documents that fulfilled any of the following criteria were excluded: 1) They exclusively consisted of a set of recommendations; 2) They contained recommendations exclusively targeted to patients; 3) They were exclusively oriented to health services organization; 4) They lacked information related to the methodology applied (the inclusion of the word “consensus” was enough for not excluding the document); 5) They were referred to as guidelines, but were undertaken by only one author without any reference to the methodology applied; 6) They were not exclusively developed by an Argentine institution; 7) The year of development or diffusion of the document was not stated.

Subsequently institutions were classified according to the level of the health system to which they belong (macro, meso and micro level). Table 2 shows the definition used to describe each level.

All CPG were included in the quality assessment (QA) phase.

A total of 30 health professionals distributed throughout the country were invited to participate in the study. All those professionals who accepted to participate were provided with the Spanish version of the AGREE instrument, the Spanish and English version of the Training Manual and an instructive document. One CPG was selected from the database and sent to all professionals for assessment. An individual feed-back was provided to each of the evaluators. Only those professionals that successfully completed this training were formally accepted as appraisers (n = 23). Demographic and professional data was collected on all appraisers.

The AGREE instrument was selected for the assessment phase, as it was the only one that has been internationally validated and has a Spanish version already tested in Spain.^{12,15} The instrument provides a systematic framework for assessing key dimensions of the quality of guidelines (Table 3).¹⁶

Guidelines were assigned taking into account the expertise and specialty of each appraiser. All appraisers returned the results of the assessment by e-mail, using an excel format (that was previously developed and reviewed through a pilot test). As the study was done without financial support, none of the appraisers received any honorarium. The results were entered in a database. Domain scores of each CPG were calculated as recommended by the AGREE Collaboration. Scores of individual items in each domain were summed and standardized as a percentage of the maximum possible score for that domain.

The internal consistency of each domain was evaluated using Cronbach's Alpha. The reliability between appraisers was determined for each question and each domain of the AGREE. Intraclass correlation coefficients (ICC) were calculated within each pair of

appraisers and across the pool of appraisers. ICCs and kappa values above 0.75 were considered to represent good, 0.40–0.75 moderate and <0.40 poor reliability.

Feasibility of the instrument was assessed through an *ad-hoc* instrument that contained two dimensions: usefulness and simplicity; both dimensions were assessed through a 1-5 scale, being 5 the highest score.

Results

Guidelines production in Argentina

A total amount of 431 pCPG were identified through the combination of both searching strategies (EDS and IS). In 361 (83%) cases, pCPG had been developed by institutions that belonged to one level of the health system; in 70 (17%) cases, multiple institutions from different levels participated in the process (Table 4). Approximately 60% of the pCPG were produced by institutions from the meso and micro level.

Most guidelines (363) were identified through IS. Websites of 247 institutions were assessed, but only 17% (42) of these institutions have posted any guidelines on their website. Only 16% (68/431) of the pCPG were published in indexed journals and identified through EDS.

Two hundred and eighty seven (65%) pCPG included recommendations related to evaluation, diagnosis, treatment or management of specific diseases. Only 15% were oriented to prevention and screening practices (Figure 1). Production of pCPG shows a constant increment, specially during the period of 2000-2004. Management pCPG remained as the most frequent category along the years. Prevention pCPG slightly increased from 1994 to 2003, but they decrease in the last year. Technology

Assessment pCPG appeared in 2001 and show a steady growth during the following years (Figure 2).

Quality Assessment

Only 33% (144) of the 431 pCPG met the QA criteria. The most frequent reasons for QA exclusion were lack of one of the following: i) year of publication or diffusion, ii) a methodology section, and iii) references or authors' identification in the case of consensus (Table 5). Of the 143 CPG selected, 101 (71%) were assessed and 43 were eliminated (Table 6).

The majority of the CPG assessed received very low scores in nearly all domains (Figure 3). Over 80% of the CPG were qualified with scores lower than 50%, except in those domains corresponding to "Clarity" and "Scope". In comparison to the results of the other domains, clarity was the best qualified aspect of CPG. Although total amount of CPG increased progressively during the study period, there was no statistically significant difference in the quality standardized scores (Figure 4).

ICC and α Cronbach for each domain were in all cases moderate or high (0.46-0.74), except for Editorial Independence which showed very low values (0.20) (Table 7). Most of appraisers considered that the AGREE instrument was useful and simply to apply (Usefulness, Median value = 5 and Simplicity, Median Value: 4). Scope was selected by the group of appraisers as the most difficult domain to be assessed. Average time (standard deviation) employed per CPG assessed was 58 (\pm 36) minutes.

Discussion

In the last years, development of guidelines in Argentina has progressively increased; however, quality did not improve. This situation could be clearly resumed in the phrase of Sudlow and Thomson: "Quantity but no quality".¹⁷

Guidelines production has been non-systematized throughout the whole period: integration of the interests, preferences and activities of multiple stakeholders in guideline development was

rarely observed. This non-integrated participation could be partially explained by the complexity of the Argentine health system but, specially, by the absence of a clear and explicit policy for guidelines production and evaluation during the period under assessment. On the whole, institutions of the three levels of the health system participated in guideline development, although reporting of the methods employed in this process was poor and need to be improved. Quality of guidelines that were assessed was distant to ideal: scores were low and very low in all domains of the AGREE instrument. Even when standards recommended by the AGREE Collaboration should be used as a reference for improving this process in the following years, two issues should be considered at the moment of interpreting these findings. In first place, standards proposed by the AGREE instrument could be relatively high for the context of a LMIC and specially if it is taken into account the fact that, except for the last two years (2003 and 2004), the period during which Argentine guidelines production was described preceded the year of diffusion of this instrument (2003). Secondly, low quality scores of Argentine guidelines could be explained by a slower penetration and consolidation of the Evidence-Based Movement in LMIC countries in comparison to developed countries. As described by Burgers, development of guidelines in Europe, Australia and North America started in the 80's and 90's.³ In the United States, the Consensus Development Program at the National Institute of Health, developed its first guideline in 1977. In the last 30 years, all these organizations have accumulated a vast experience in guideline development, dissemination and implementation. Currently, principles of Evidence-Based-Medicine dominate almost all of these national guideline programs. The creation of international networks, like the GIN, as well as the conduct of projects like the AGREE, have clearly contributed to the improvement and standardization of these processes in the participant countries. Contrastingly, LMIC countries, with a few exceptions, did not take part of these experiences. Diffusion and dissemination of appropriate methods for evidence-based guidelines development is limited in these countries. According to the results of this study, in Argentina, as late as 2004, this process was not systematized and still relied heavily on the opinion of experts.

Lack of systematization in guidelines production as well as low quality scores were also reported by other authors, even from developed countries. In Finland, 719 guidelines were produced since 1989 to 1995: the structure and quality of evidence supporting guidelines recommendations were variable and only two were based on meta-analysis¹⁸. Similarly, Graham et al. assessed 217 drug guidelines from 1994 to 1998: quality varied significantly by developer, publication status and drug company sponsorship and no substantial improvement in guideline quality was observed over the 5-year study period⁹. Many other studies mainly performed in developed countries showed similar findings.^{6-8, 10-13} There is only one CPG quality assessment in Latinoamerica that precedes the Argentine research, which was performed in Brazil. In this study, twenty-eight guidelines developed by the Brazilian Medical Association were independently assessed by 2 appraisers using the AGREE instrument. In general, scores are higher than those observed in the Argentine research, although it should be taken into account that guidelines included in the Brazilian experience belonged to only one institution, which has an explicit policy to develop evidence-based guidelines.¹⁴

Even when scores in all domains were low in the case of Argentine guidelines, the three most critical were rigor of development, applicability and editorial independence.

Respecting to rigor of development, description of the search strategy as well as inclusion criteria applied in the selection of the most relevant and pertinent evidence were practically inexistent. Lack of bibliographic references was one of the most frequent reasons for excluding pCPG from quality assessment and in those CPG that were assessed, the link between evidence and recommendations was not described except for a few cases (4 guidelines). Even when currently there is broad agreement on the need for systematic reviews to inform recommendations, this type of evidence was rarely referred in Argentine guidelines.¹⁹

Language barriers and limited accessibility to updated biomedical literature could have contributed to this situation. In this sense, the role of the Argentine Cochrane Centers might be relevant in the future, specially by promoting the use among guideline producers of

resources like the Cochrane Library Plus, which contains the Spanish version of systematic reviews produced by the Cochrane Collaboration.²⁰ As it was reported by Varonena and col., cooperation between Cochrane Centres and Guideline Development Organizations showed to be very positive in many senses.²¹

Another area that should be improved in the following years is related to the applicability domain. In many cases, recommendations were literally taken from international guidelines, without addressing potential barriers to their implementation in the context of the Argentine health system. Since the cost of producing evidence-based guidelines is relatively too high for health budgets of LMIC countries, a systematic methodology to adapt international guidelines would be an efficient way of improving not only the quantity but also their quality.²² Internationally developed guidelines can be adapted to the national context, representing a considerable saving of money. However, an explicit and systematic adaptation process should be performed as guidelines' applicability and transferability can be strongly influenced by different factors, like population needs (prevalence of disease, baseline risk status), setting (availability of resources) and other factors that modify translation of recommendations into practice.²³

Finally, practically none Argentine guideline reported conflict of interests or funding sources. Lack of transparency was also reported by Papanikolaou et al. in an evaluation of 191 published guidelines: only 7 (3.7%) disclosed potential conflicts of interest.²⁴ In the case of Argentine guidelines, omission to declare conflicts of interest could have been due either to lack of disclosure by guideline authors or to lack of formal declaration processes by institutions that endorsed the guideline. Omission could have been unintentional or, on the contrary, intentional (financial ties might have existed in some situations and deliberately hidden by guideline authors). However, regardless of the intent of guideline developers' actions, *explicit* declaration of conflict of interests *at the beginning of the process* is strongly recommended by most international organizations as a way of reducing the probability of biased recommendations and increasing guidelines' credibility.²⁵

The AGREE instrument was selected amongst other quality assessment instruments as it is the only one that covers practically all the relevant dimensions of the guideline development process; it has fewer items and uses a numerical scale that facilitates the analysis and comparison of the results.^{26,27} The AGREE instrument has been widely used all over the world, mainly as a result of its translation into many other languages, including Spanish. There is scarce evidence related to the application of the instrument spanish version. In 2005, Navarro MA et al reported the results of a quality assessment on 61 guidelines developed in Spain. Like in the Argentine experience, the instrument showed to be feasible to apply and reliability was satisfactory. Appraisers of both research found the instrument useful and simply.¹²

There are a number of methodological issues that should be addressed. First of all, evaluation was restricted to guidelines that were diffused and identified on Websites or in journals. Diffusion is not the same as development as there might have been guidelines produced and used in health institutions that could have not been identified by the searching strategies applied in this research. In spite of this limitation, the study was focalized on those guidelines diffused by well-known and reputed institutions in Argentina, which have a high probability of being adopted by healthcare professionals. Secondly, Internet searching was not exhaustive at the meso level, concretely in hospitals: as a reduced number of these institutions have Websites, only 3 out of 10 eligible hospitals could be assessed. Thirdly, even when the AGREE collaboration strongly suggests 4 appraisers per each CPG, this could not be performed because of lack of resources. All researchers and appraisers work add-honorem. However, in spite of this fact of including only two appraisers per guideline, reliability scores were acceptable. In fourth place, only CPG documents were considered for the assessment. Another important recommendation of the AGREE Collaboration is to retrieve any relevant information related to the guideline development process at the time of its evaluation. Except from two or three institutions that include a full description of their CPG programs on the websites, the guideline document was the only one available. However, it should be taken into account, as Graham highlighted in his article, that these documents are also the only one

available for practitioners at the moment of making a clinical decision. In all likelihood, they do not have any additional information that the one that was provided to appraisers. Finally, there are some limitations inherent to the instrument applied. Quality of guideline is assessed on the basis of what is reported: quality of reporting is not the same as quality of the development process. As in other quality assessment studies, no content analysis of the recommendations was performed.

To our knowledge this is the first time a study of this kind has been undertaken in Argentina, and, except from the Brazilian research, in Latinoamerica. Its execution was the first step in the building of a network of professionals interested in improving CPG development, dissemination and implementation in the country. Its findings might be very useful in the set up of a national evidence based guideline development program.

Conclusions

This study was one of the formers that sistematically employed the AGREE instrument for the critical assessment of guidelines produced in a LMIC. The AGREE instrument can serve as a model to identify improvement opportunities in the guidelines development process of these countries. In this sense, this research shows the low quality of guidelines produced and points out areas to which training iniatitives should be oriented.

Guideline development and diffusion in Argentina from 1994 to 2004 shows a constant increment, although quality of reporting did not improved; moreover, in some aspects it seemed to decline. Institutions involvement in this process was dispersed, rarely integrated, and not systematized. A national debate between main stakeholders is urgently needed in order to contribute to the definition of a clear and explicit policy for CPG development, dissemination and implementation in the country.

Competing interests

All the authors expressed that there are none competing financial or non-financial interests.

Authors' contributions

MEE conceived the study, designed the protocol and coordination of the research, performed the Internet Search, registered pCPG in the database, selected the CPG from the database, was in charged of appraisers training and appraised guidelines, performed the statistical analysis, interpreted the data, drafted the manuscript. ZO participated in the design of the protocol and coordination of the research, selected the CPG from the database, interpreted the data and helped to draft the manuscript. MGD performed the database electronic searching, appraised guidelines, interpreted the data, and helped to draft the manuscript. ECh, RM and RB appraised guidelines, interpreted the data and helped to draft the manuscript.

Acknowledgements

The authors would like to express their special thanks to all the health professionals that agreed to participate as guideline appraisers: Betina Nishishinya; Elsa Andina; Lydia Otheguy; Martin Olmos; Mónica Confalone; Graciela Berenstein; Gustavo Badarotti; Martin Díaz; Florencia Koch; Irma Aznar; Mariana Pasquali; Graciela Sarmiento; Matías Mirofsky; Gustavo Piñero; Marçia del Carmen Esandi; Noberto Schapochnik; Martín Urtasun; Facundo Goya.

They would also like to express their gratitude to María de los Angeles Lopez-Olivo, for her invaluable feedback and input and Dr. Jeremy Grimshaw for his support to this research.

Figures

Figure 1 - Distribution of guidelines according to IOM categories

Note: () 3 pCPG could not be classified in any of this three categories.*

Figure 2 - Temporal distribution of the different types of guidelines

Figure 3- Distribution of CPG per each domain of the AGREE

Figure 4. Temporal evolution of the median standardized score for each domain

Tables

Table 1 – Electronic Database Strategy employed in Medline under PubMed Platform

Table 2. Criteria applied to define macro, meso and micro level health institutions eligible for Internet Searching.

Table 3. Domains of the AGREE instrument

Table 4. Distribution of pCPG by level of the health system and developer institution.

Table 5. Reasons for exclusion from QA

Note: () Some pCPG met more than one criteria.*

Table 6. Reasons for elimination from QA

Table 7. Reliability Scores of the AGREE instrument

References

1. **Clinical Practice Guidelines: Directions for a New Program.** Edited by Field MJ; Lohr L. **Committee to Advise the Public Health Service on Clinical Practice Guidelines**, Institute of Medicine; 1990: 19-32.
2. Woolf SH; Grol R; Hutchinson A; Eccles M; Grimshaw J. **Clinical guidelines: Potential benefits, limitations, and harms of clinical guidelines.** BMJ 1999; 318:327-30.
3. Burgers J; Grol R; Klazinga N; Makela M; Zaat J, for the AGREE Collaboration. **Towards Evidence-Based Clinical Practices. An international survey of 18 clinical guideline programs.** International Journal for Quality in Health Care; 2003; 15 (1): 31-45.
4. **Guidelines International Network (GIN).** Available in: [<http://www.g-i-n.net>].
5. **The AGREE Collaboration. Appraisal of Guidelines, Research and Evaluation.** Available in: [<http://www.agreecollaboration.org/>].
6. **The AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project.** Qual Saf Health Care 2003; 12:18-23.
7. Shaneyfelt T; Mayo-Smith M; Rothwangl J. **Are Guidelines following guidelines? The methodological quality of Clinical Practice Guidelines in the peer-review medical literature.** Jama 1999; 285: 1900-1905.
8. Cluzeau F; Littlejohns P; Grimshaw J; Feder G; Moran S. **Development and application of a generic methodology to assess the quality of clinical**

-
- guidelines.** International Journal for Quality in Health Care; 1999; 11 (1): 21-28.
9. Grilli R; Magrin N; Penna A; Mura G; Liberati A. **Practice guidelines developed by specialty societies: the need for a critical appraisal.** Lancet 2000; 355: 103–06.
10. Graham I; Beardall S; Carter A; Glennie J; Hebert P; Tetroe J; McAlister F; Visentin S; Anderson G. **What is the quality of drug therapy in Canada?** CMAJ 2001; 165 (2):157-63.
11. Harpole L; Kelle M. Schreiber G; Toloza E; Kolimaga J; McCrory D. **Assessment of the Scope and Quality of Clinical Practice Guidelines in Lung Cancer.** CHEST 2003; Suppl 123:7–20.
12. Navarro Puerto MA; Romero FR; Dominguez AR; Gutiérrez Ibarlucea I; Gago H; Ortiz del Río CA; Iturrioz RR; Batarrita JA. **¿Las Guías que nos guían son fiables? Evaluación de las Guías de Práctica Clínica Españolas.** Rev Clín Esp 2005; 205(11):533-40.
13. MacDermid JC; Brooks D; Solway S; Switzer-McIntyre S; Brosseau L; Graham I. **Reliability and validity of the AGREE instrument used by physical therapists in assessment of clinical practice guidelines.** BMC Health Services Research 2005, 5:18 doi:10.1186/1472-6963-5-18.
14. Nobre MR; Bernardo WM; Jatene FB. **Critical Appraisal of 28 Guidelines developed by the Brazilian Medical Association. In Proceedings of the Second International Conference of Evidence Based-Health Care Teachers & Developers:** 10-14 September 2003; Palermo (Italy). Available in: [http://www.ebhc.org/2003/abstracts_book.pdf]. (Last accessed: July 2007).

-
15. **The AGREE Instrument Spanish version.** The AGREE Collaboration. Available in: [http://www.agreecollaboration.org/translations/index.php?lang=es]. (Last accessed: July 2007).
16. **Appraisal of Guidelines for Research & Evaluation Instrument Training Manual.** The AGREE Collaboration, January 2003. Available in: [http://www.agreecollaboration.org/pdf/aitraining.pdf].
17. Sudlow M; Thomson R. **Clinical guidelines: quantity without quality.** Qual Health Care 1997; 6: 60-61.
19. Varonen H; Makela M. **Practice guidelines in Finland: availability and quality.** Quality in Health Care 1997; 6: 75-79.
19. Cook D; Greengold NL; Ellrodt G; Weingarten S. **The Relation between Systematic Reviews and Practice Guidelines.** Ann Intern Med 1997; 127 (39:210-216).
20. Cochrane Library Plus. Available in: [http://www.update-software.com/Clibplus/ClibPlus.asp]. (Last accessed: July 2007).
21. Varonena H; Jousimaab J; Helin-Salmivaarac A; Kunnamoia I. **Electronic primary care guidelines with links to Cochrane reviews—EBM Guidelines.** Family Practice 2005; 22: 465–469.
22. Graham ID; Harrison MB; Brouwers M; Davies BL; Dunn S. **Facilitating the Use of Evidence in Practice: Evaluating and Adapting Clinical Practice Guidelines for Local Use by Health Care Organizations.** JOGNN, 31, 599–611; 2002.

-
23. Schünemann HJ, Oxman AD, Fretheim A: **Improving the Use of Research Evidence in Guideline Development: 13. Adaptation, applicability and transferability.** Health Res Policy Syst 2006.
24. Papanikolaou GN, Baltogianni MS, Contopoulos-Ioannidis DS, Haidich AB, Giannakakis IA, Ioannidis JPA. **Reporting of conflicts of interest in guidelines of preventive and therapeutic interventions.** BMC Medical Research Methodology 2001., 1(3). Available at: [<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=11405896>] (Last accessed: January 2007).
25. Boyd E, Bero L. **Improving the Use of Research Evidence in Guideline Development: 4. Managing conflicts of interest.** Health Res Policy Syst 2006. Available at: <http://www.health-policy-systems.com/content/4/1/16> (Last accessed: January 2007).
26. Rico Iturrioz R; Ibarluzea-Gutierrez I; Batarrita JA; Navarro Puerto A; Reyes Domínguez A; Marín León I; Briones Perez de la Blanca E. **Valoración de Escalas y Criterios para la evaluación de Guías de Práctica Clínica.** Rev Esp Salud Pública 2004; 78: 447-457.
27. Vlayen J; Aertgeerts B; Hannes K; Sermeus W; Ramekers D. et al. **A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit.** Qual Saf Health Care 2005; 17:235-242.

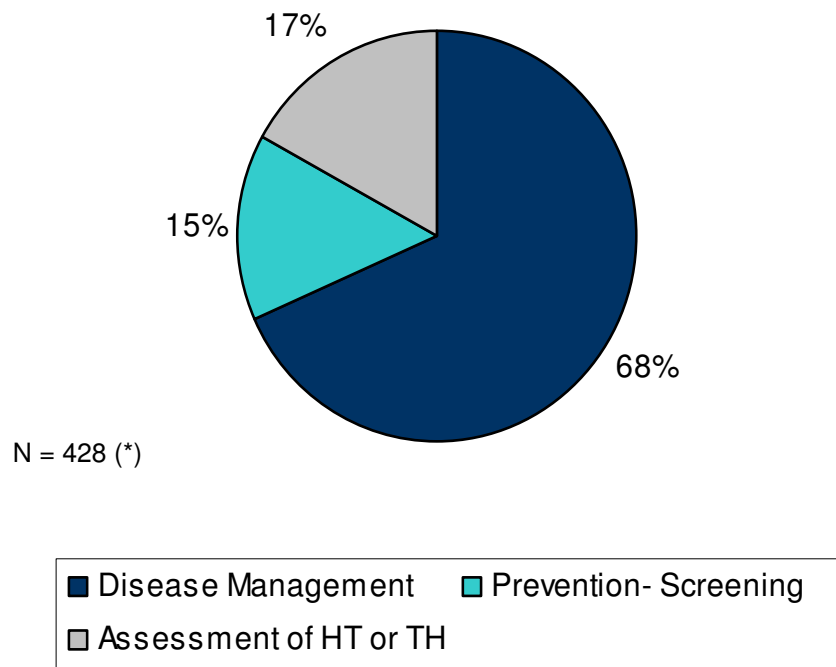


Figure 1

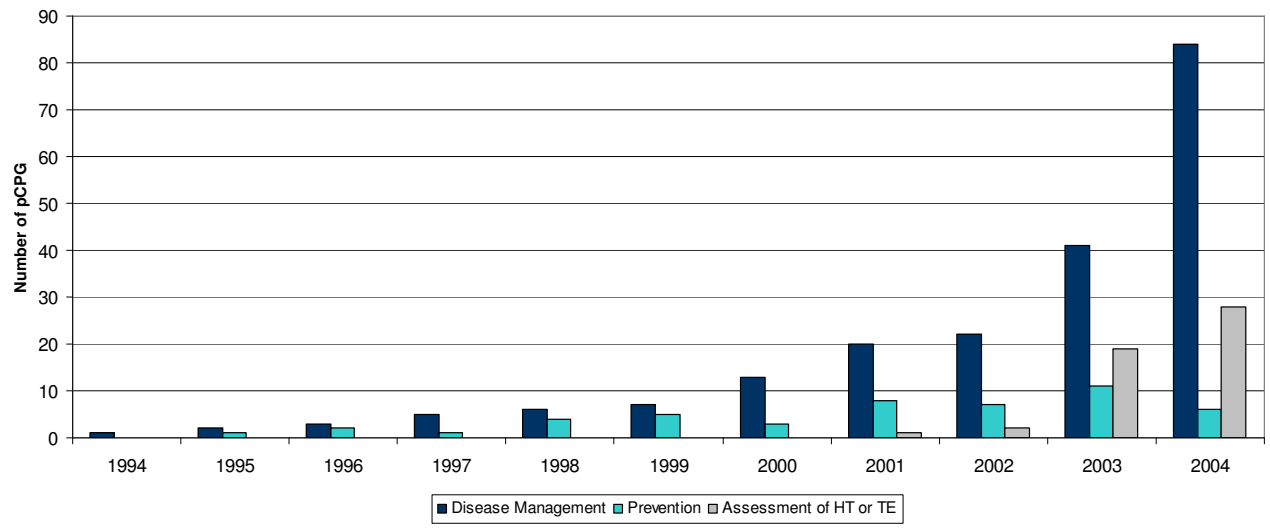


Figure 2

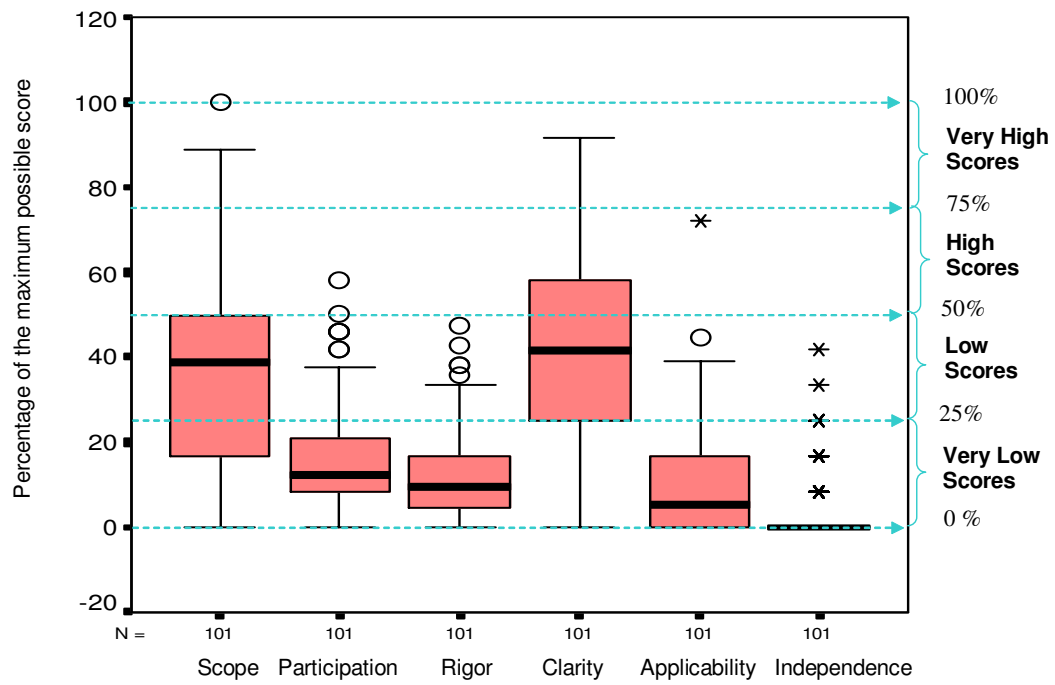


Figure 3

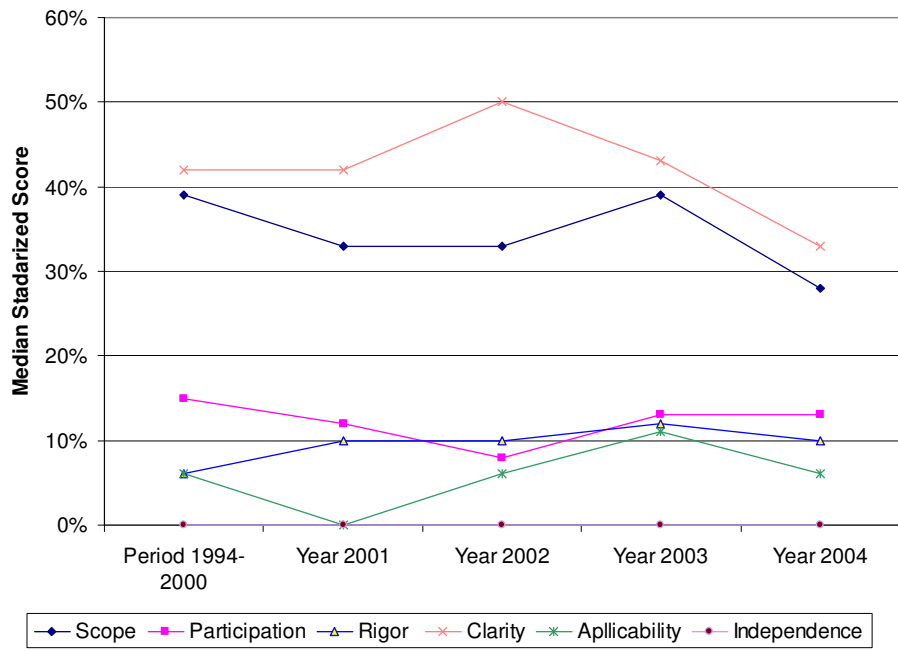


Figure 4

Additional files provided with this submission:

Additional file 1: table 1.doc, 26K

<http://www.implementationscience.com/imedia/2121833178182394/supp1.doc>

Additional file 2: table 2.doc, 26K

<http://www.implementationscience.com/imedia/8405236421823944/supp2.doc>

Additional file 3: table 3.doc, 27K

<http://www.implementationscience.com/imedia/4464782091823944/supp3.doc>

Additional file 4: table 4.doc, 35K

<http://www.implementationscience.com/imedia/4859967421823944/supp4.doc>

Additional file 5: table 5.doc, 26K

<http://www.implementationscience.com/imedia/1728233829182394/supp5.doc>

Additional file 6: table 6.doc, 24K

<http://www.implementationscience.com/imedia/1140885364182394/supp6.doc>

Additional file 7: table 7.doc, 28K

<http://www.implementationscience.com/imedia/6819552151823944/supp7.doc>