Locator web
Article 44.098

Characteristics of primary care clinical guidelines associated with greater structural quality

J. Saura-Llamas[a], P.J. Saturno Hernández[b], J.R. Romero Román[a], J.M. Gaona Ramón[c], J.J. Gascón Cánovas[d] and Grupo de Evaluación y Mejora de los Protocolos Clínicos [Clinical Guideline Evaluation and Improvement Group][e]

Aim. To identify characteristics associated with greater structural quality of clinical guidelines.
Design. Cross-sectional study.
Setting. Health centers in the region of Murcia (southeastern Spain).
Main outcome measures. All clinical practice guidelines and protocols developed between January 1985 and January 1994 were reviewed. Of the 470 documents originally obtained, 462 were evaluated and 8 were excluded because of missing data. The quality of document design was evaluated in all materials. The rate of criteria compliance was calculated for each document. The characteristics that were associated with protocol quality were identified in two types of multivariate analysis: multiple regression (with compliance rate as the dependent variable) and logistic regression (with compliance rate referred to the mean as the dependent variable).
Results. Both analyses showed that structural quality was associated with specific health care areas, multidisciplinary design ($p < 0.001$), reference to chronic health problems ($p < 0.001$), design of the document specifically as a clinical practice guideline ($p < 0.001$), and reference to the health services offered at a given center ($p < 0.001$). In some analyses, greater quality appeared to be associated with heath centers that were also teaching centers, reference in the document to health care, and women´s health programs.
Conclusions. Document quality varied significantly in different health care areas, and certain characteristics (chronic health problems, multidisciplinary design and specific design, reference to specific health services offered) were associated with greater document quality. Reference to acute health problems, design by only one type of professional (physicians or nurses), inclusion as part of a larger program, and lack of reference to specific health services offered at a given center were characteristics with a greater risk for low document quality.

Key words: Primary care. Quality of care. Clinical practice guidelines. Protocols.

CARACTERÍSTICAS DE LAS GUÍAS CLÍNICAS DE ATENCIÓN PRIMARIA QUE SE ASOCIAN A UNA MAYOR CALIDAD ESTRUCTURAL

Objetivo. Identificar las características que se asocian a una mayor calidad estructural de las guías clínicas.
Diseño. Evaluación transversal.
Emplazamiento. Centros de salud de la región de Murcia.
Participantes. Documentos (y profesionales) de esos centros.
Mediciones principales. Son objeto de estudio todas las guías de práctica o protocolos elaborados de enero de 1985 a enero de 1994, obteniéndose 470, de las que se evalúan 462 (se rechazan 8 por falta de datos). Se valora la calidad del diseño de los protocolos. Se calcula la ratio de cumplimiento de criterios para cada documento. Se identifican las características que se asocian a la calidad de los protocolos con dos análisis multivariantes: regresión múltiple (variable dependiente la ratio cumplimiento) y regresión logística (variable dependiente la ratio de cumplimiento en relación a la media).
Resultados. En ambos análisis una mayor calidad estructural se asocia con una determinada área de salud, elaborados de manera multidisciplinaria ($p < 0,001$), referidos a un problema de salud crónico ($p < 0,001$), elaborados específicamente como tales ($p < 0,001$) y relacionados con la cartera de servicios ($p < 0,001$). En alguno de los análisis parece asociarse una mejor calidad con que el centro de salud sea docente, que la guía se refiera a la asistencia y del programa de salud de la mujer.
Conclusiones. La calidad de los documentos varía significativamente según el área de salud, y determinadas características (problemas de salud crónicos, elaboración multidisciplinaria y específica, y relación con la cartera de servicios) se asocian a una superior calidad de los documentos.
Las características de problemas agudos, elaboración uniprofesional, ser parte de un programa y no relacionados con la cartera de servicios se mostraron como de mayor riesgo para una baja calidad.

Palabras clave: Atención primaria. Calidad asistencial. Guías de práctica clínica. Protocolos.

Spanish version available at
www.atencionprimaria.com/43.752

A commentary follow this article (pág. 532)

[a]Specialist in Family and Community Medicine, Barrio del Carmen Health Center, Murcia.
[b]Titular Professor of Preventive Medicine and Public Health, University of Murcia.
[c]Pediatrician, Torre Pacheco Health Center, Murcia.
[d]Adjunct Professor of Preventive Medicine and Public Health, University of Murcia.
[e]Grupo de Evaluación y Mejora de los Protocolos Clínicos: C. Quirós Bauset, R. Gomis Cebrián, B. Lor Esteban, I. Medina, F. Miralpeix and M. González Berberá.

Correspondence:
José Saura Llamas, C/ Atenas 21, 30120 El Palmar (Murcia), Spain.
E-mail: csgoya@iname.es

This research was carried out with financial support from FIS project 94/1187. None of the authors has any conflict of interest.

The data in this article were first reported in a doctoral dissertation supported by FIS project 94/1187. They were communicated at the II Congreso de Calidad Asistencial [Second Congress on Quality of Care] held in Murcia in March 1999.
Manuscript accepted for publication 4 July 2001.

Introduction

Good design and the prevention of quality-related problems through planning are two necessary components of programs for health care quality.[1-3] The development of clinical practice guidelines is one technique that can be used to design quality of care. These documents can take a variety of forms, as noted in appendix B of the classic IOM report,[4] which gives examples of different presentation formats for algorithms, clinical protocols, and other documents. Among the different working techniques clinicians use, the development of clinical practice guidelines for specific medical problems is probably one of the most widely known and often used strategies, as it is an essential part of the process of clinical care improvement.[5-7] Thus the development of such guidelines is one of the most strongly supported activities at the international level as part of programs aimed at managing health care,[8-10] and as an important field of research.[11-14]

Clinical practice guidelines are tools that facilitate decision-making by physicians and that help them to fight against uncertainty and diminish variability in clinical practice. However, for the guidelines to be effective they must be satisfy minimum requirements for formal quality, i.e., they must be validly structured so that the results of their application can be evaluated. A bad tool can simply be ignored, or it may invalidate the clinical results obtained as a result of its implementation. As part of the ongoing research by our working group,[15-19] and as a follow-on to our analysis of the structural quality of these documents,[19] we set out in the present study to identify the variables that influenced formal quality of clinical practice guidelines.

Material and methods

According to the 1998 census, the autonomous region of Murcia in Southeastern Spain has 1115068 inhabitants. Most primary health care in this region is provided through the public network of health care centers; the present study was based on documents from these centers.

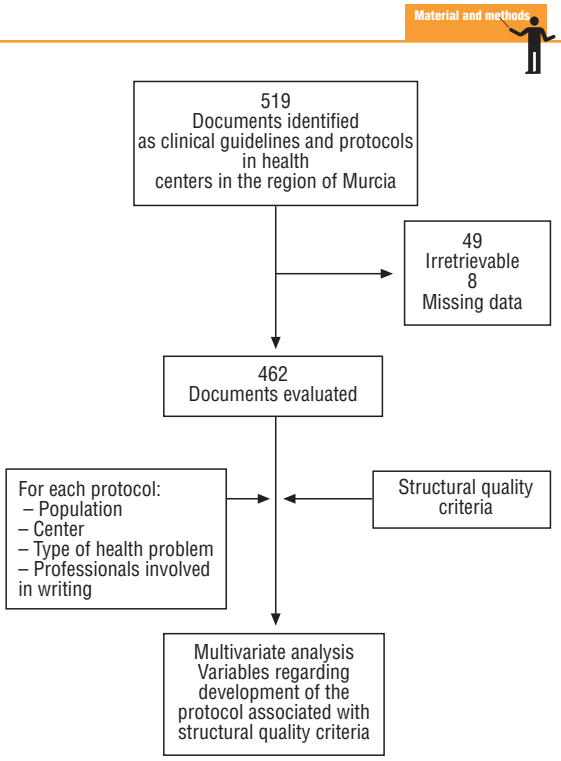
On the basis of an earlier study at one health center,[16] our research group undertook a retrospective study of all protocols developed in the region of Murcia between January 1985 and January 1994. During that period 31 of the 39 existing health centers used formal protocols for at least some of their activities.

Protocols were identified by searching the regional registers of the Consejería de Sanidad (health council of the regional government) and the «gerencias de atención primaria» (primary care management offices), and by a telephone survey of all health center coordinators. A total of 470 clinical protocols were submitted to the research group or obtained in the course of a visit to the center. Eight documents were excluded from the analysis because of missing data (authors, health center, etc.); in all, 462 documents were evaluated. The independent variables and their categories are shown in table 1.

The *selection of structural quality criteria* and the description of each criterion are detailed along with other characteristics of the materials and methods in an earlier publication.[19] To *evaluate structural quality of the protocols* we used compliance with each of the 9 quality criteria established for each protocol.

To identify those characteristics that had an overall influence on structural quality of the documents we investigated the relation between the independent variables and two dependent variables: compliance rate and ratlog. *Compliance rate* for a given protocol is the relationship between the number of structural quality criteria this document met and the total number of structural criteria considered (9). The variable designated *ratlog* was a qualitative variable with two categories: 0 if the percentage of compliance with individual criteria was higher than the mean compliance rate for all protocols, or 1 if it was lower.

Most of the independent variables were qualitative in nature. The *dichotomous* variables were: whether the protocol formed part of a larger health care program, accreditation of the protocol, and whether it dealt with a health service that formed part of the services normally provided by the center. Qualitative variables with more than two categories were: health care area, teaching accreditation of the center, participation of more than one type of professional in the development of the protocol, population group served by the protocol, nature of the health problem, and type of activity covered by the protocol. The variables «year



General scheme of the study

Cross-sectional study of the application of quality criteria for clinical guidelines, and evaluation of the variables associated with compliance with criteria for structural quality.

TABLE 1 **Independent variables and their categories**

Independent variable	Category
Health care area Murcia	Cartagena
	Lorca
	Caravaca
	Jumilla-Yecla
	Cieza–Molina
Year health care center was opened	1985-1987
	1988-1990
	1991-1993
Teaching accreditation	Undergraduate and graduate
	Graduate onlyo
	Not accredited
Year of accreditation	Year
Year protocol was written	Year
Professionals involved in writing	Physicians
	Nurses
	Other
	Multidisciplinary
Population served	Children
	Women
	Adults
	Elderly
	Other
Type of activity	Health education
	Acute care
	Follow-up care
	Rehabilitation
	Various
	Other
Type of health problem	Acute
	Chronic
	Various
	Other
Included in services normally offered by the center	Yes
	No
Included in a health care program	Yes
	No
Accredited	Yes
	No

the center was opened» and «year of accreditation of the protocol» were also treated as categorical variables with four possible «values»: 1986, 1987, 1992 or 1993.

Nonbinary qualitative variables with k categories were converted to dummy variables and broken down into k-1 imaginary binary variables such that the non-introduced variable served as the reference category.[20]

Multivariate analysis was used to search for characteristics of the protocols or of their writing and development process that were significantly associated with greater quality of the protocol design. Two types of analysis were used: *multiple regression* and *logistic regression*.

Multiple regression

Multiple regression analysis was used to try to identify those characteristics that improved the protocols regardless of whether the compliance rate was below the mean for all protocols analyzed. That the data fulfilled the requirements for using this approach was verified (linear relation, normal distribution, homoscedasticity and absence of colinearity) with analysis of residuals and analysis of tolerance.[21] The «enter» method was used with compliance rate (mean compliance per protocol) as the dependent variable. As independent variables we used those mentioned above. Year of accreditation of the protocol was subsequently excluded because of the low number of documents (175 out of 462) for which this information was available.

Logistic regression

Logistic regression was used to try to identify the characteristics of the best protocols, i.e., to find out which features were associated with a level of structural quality higher than the mean for all protocols.

We calculated the adjusted odds ratio (OR) at a 95% confidence interval for each characteristic analyzed as an independent variable and for the variable «ratlog».

All statistical analyses were done with the PC(+)#r Statistical Package for Social Sciences (SPSS).

Results

Of the 519 documents identified from different registries, we obtained 470 (49 were irretrievable); 462 of these underwent full evaluation.

Only a few documents had missing data for the variables «year the center was opened» (5 cases, 1.1% of the total sample) or «inclusion in services offered by the center» (45 cases, 9.7%). We considered these rates to be low and within our expectations. Many protocols did not specify the «year the protocol was written» (311 cases, 67.3%) or the «year of accreditation of the protocol» (175 cases, 37.9%), as this information was missing from the document itself and from the registry. Therefore these two variables were excluded in most of the subsequent analyses.

Multiple regression with the mean number of defects per document as the dependent variable showed that better document quality was significantly associated with women´s health programs ($p<0.03$) and inclusion in the services normally offered by the center ($p<0.001$). Features that were associated with worse document quality were health care area other than that those corresponding to the city of Murcia (which was used as the reference location), non-teaching health center, protocol written by nurses only or by physicians only, protocol dealing with an acute

TABLE 2 **Variables that significantly influenced quality in the multiple regression analysis. Compliance rates (β) per variable for the entire sample of documents analyzed**

Variable	Category	β	significance
Health area (reference: Murcia)	Cartagena	–0,16433	0,0000
	Lorca–Caravaca-Jumilla-Yecla	–0,01727	ns
	Cieza-Molina	–0,05586	0,0388
Teaching accreditation (undergraduate and graduate)	Graduate	–0,03168	ns
	Nonteaching	–0,06129	0,0024
Professionals (multidisciplinary)	Physicians	–0,15052	0,0000
	Nurses	–0,08164	0,0005
Population group (adult)	Children	0,06965	ns
	Women	0,06355	0,0321
	Elderly	6,3568	ns
	Other	–0,02051	ns
Type of health problem (chronic)	Acute	–0,06604	0,0200
	Other	–0,08073	0,0007
Type of activity (acute care)	Education and prevention	–0,04563	ns
	Follow-up care	–0,10612	0,0002
	Other	–6,788	ns
Part of a program	Yes	–0,12846	0,0000
Accredited	Not accredited	–0,03518	ns
Regular services	Yes	0,07280	0,0011

health problem, and inclusion of the protocol as an annex to a larger health program rather than being developed specifically as an independent protocol or clinical guideline (table 2).

Table 3 (showing characteristics that yielded significant results) details the results of the logistic regression analysis with the dependent variable «ratlog», which identified documents whose percent compliance rate was greater than or equal to the mean. Table 4 (showing ORs and 95% confidence intervals) summarizes the results of the overall analysis of protocol quality. The ORs showed significant associations for the Cartagena and Cieza-Molina health care areas in comparison to the Murcia area. Other significant associations were identified for protocols written by physicians alone in comparison to protocols written by a multidisciplinary team, protocols designed for acute or other health problems in comparison to chronic health problems, protocols designed for follow-up procedures rather than for acute care, protocols that formed part of a health care program, and protocols designed for a service normally offered by the health center. The likelihood of a given clinical guideline being of greater than average quality was at least threefold as high (OR=2.98) if the document dealt with a service normally offered by the center than if it was developed for a service not normally provided by that center. Document quality was worse for protocols included as an annex to a larger program in comparison to those that were developed specifically as an independent document.

Discussion

This report details our experience in an evaluation of structural quality of protocols or clinical guidelines developed by primary care teams in the region of Murcia during a relatively long (8-year) period.

Limitations of the design

The design and application of the method used here were straightforward and have been evaluated and validated previously.[1-3] The approach has been shown to be feasible and useful in primary health care, and its external validity and reproducibility in other autonomous communities have been verified in part. The quality dimension this study investigated was scientific and technical quality.[1-3]

We evaluated nearly all documents (462) in the universe of study (519), which guarantees the representativeness of our results. The process used to search for and locate relevant documents highlighted some difficulties in obtaining copies because of the lack of a central register for the entire region or of any other type of institutional register at the time of study in at least three of the six regional health

TABLE 3 **Characteristics that influenced structural quality of the protocol for each criterion separately and for all criteria. Differences that were significant are shown (n = 462)**

		C-1	C-2	C-3	C-4	C-5	C-6	C-7	C-8	C-9	Ratio
Health area (Murcia)	Cartagena	0,26[a]		0,15[b]	0,34[a]		9,19[b]	0,01b		0,04[b]	0,19[b]
Lorca, etc						46,6[b]	0,09[b]		0,04[b]		
Cieza-Molina		0,37[a]					0,09b	2,89a	0,08b	0,30b	
Year center was opened	88-90	0,33[a]				2,89[b]			0,35[b]		
90-93	0,31[a]										
Acred. doc. pre, post.	Graduate.							5,09[a]		12,62[b]	
Nonteaching								0,29[b]			
Professionals (multidisciplinary)	Physicians	0,17[b]	0,13[b]	0,21[b]		0,19[b]			0,23[b]		0,14[b]
Nurses.		0,27[b]	0,17[b]		0,41[b]	7,51[b]	0,25[a]				
Population group (adulte)	Children		4,60[b]								
Women		4,30[b]		3,42[a]							
Elderly											
Other											
Type of health problem (chronic)	Acute	0,28[b]	0,19[b]				4,62[a]		0,40[a]		0,20[b]
Other										0,40[a]	
Type of activity (acute care)	Education and prevention			4,99[b]	0,16b					0,11[a]	
Follow-up care		0,17[b]		0,05[b]	0,26[a]			0,20[b]		0,19[b]	
Other			7,41[b]	0,16[b]							
Part of a program	Yes		0,42[b]	0,25[b]	0,25[b]	0,49[a]				0,24[b]	0,24[b]
Accredited	Not accredited							0,11b		0,16b	
Regular services	Yes.	3,84[b]						5,58[a]	2,02[a]	4,16[a]	2,98[b]
Structural quality criteria used:											

C-1, fulfils definition of a protocol according to the definition used most widely used at that time in primary care. C-2, mechanism for data recording. C-3, mechanism for evaluating quality of the protocol, at least for structural features. C-4, each protocol contains at least one algorithm. C-5, includes form to record clinical information or physical examination. C-6, maximum length 20 A-4 pages. C-7, table of contents with page numbers. C-8 no formal defects, all pages and copies legible, all pages numbered. C-9, bibliography.
[a]p < 0,05; [b]p < 0,001.

care districts. Individual health centers did not keep central registers or files of all such documents; this fact accounts for some of the missing documents, suggests that professionals may not be able to access these instruments, and identifies an area in need of improvement. When this research project was in the design stage, we expected to find about 100 documents; the final figure was nearly fivefold as high. This high number may be interpreted to reflect the positive attitude of professionals toward the development of written instruments for health care, at least at the beginning of and throughout the study period.

Construction and content validity with relation to the choice of quality criteria was ensured by the fact that the study was based on previously published work, and by the process we used to guarantee reliability (interobserver agreement).

We believe some of the possible sources of bias in data collection were obviated by the detailed design of the data collection sheet, the simplicity of the data collection process, previous training and analysis of the reliability of the observers´ criteria. Because the number of lost data was low and because they were distributed in a near-random manner across all health centers, we believe these losses did not bias our results.

Comparison with earlier studies

A search of the literature found no similar studies in Spain, thus comparisons with earlier results were difficult. One earlier study[16] yielded overall results similar to ours: an initial evaluation of one health center showed overall quality of the protocols developed there to be very low, but did not identify the factors that detracted from quality. Other documents developed at primary care centers within the framework of regular training activities such as the national Continuing Medical Education Program have been found to have higher structural quality, although they were also in need of substantial improvement in some areas.[18]

The few available studies from other countries[22-24] were only partially comparable with the present report, be-

cause of differences in document collection, in the features analyzed in the documents, in the criteria used (despite methodological similarities), or in the way the results were presented. An evaluation of 855 clinical guidelines developed at 22 health centers in the Cambridge area between 1989 and 1997 showed that 75% of the guidelines referred to clinical or disease management activities,[22] a finding that appears to be similar to our results.

Two earlier studies investigated many more quality criteria, which numbered 37 in one case[23] and 25 in another.[24] Moreover, these criteria were often much more stringent than ours. These studies also found that overall quality of the guidelines was poor: in both cases mean percentage compliance with criteria (in three groups of dimensions) was below 50% for almost all criteria. The evaluation of centers in the Cambridge area emphasized that 38% of the documents failed to state the date when they were written,[22] a problem that was much more serious in our study (67% of the documents were missing this information). In another study[23] the year of publication was given in 12.8% to 23.0% of the documents. The importance of this finding resides in the speed with which scientific evidence becomes outdated, and the consequent need to indicate when the contents of the guidelines were updated. This information, according to one review, was given in only 14.3% of all protocols.[24]

Problems with the references were found in up to 90% of all documents in the Cambridge study[22], and in 74.2% of all guidelines in the review mentioned above.[24] In our material this figure (85.1%) fell between these two percentages. Thus errors in the references appear to be one of the most important and frequent defects in primary care protocols. However, none of these studies analyzed the results with an aim to identifying the factors that were related with greater structural quality of the protocols.

Practical applicability of the results

We believe that research on the quality of health care instruments is timely and valuable because of its repercussions for (among other areas) health and the costs of health care. Our finding that the date the health center opened was not significantly related with document quality can probably be taken to imply that no improvement in quality can be expected in centers where the staff members have more experience working together. The findings allowed us to identify the factors that have a positive influence on quality. Further study would allow us to determine the conditions under which these factors occur, and might make it possible to reproduce these conditions in order to favor the production of high-quality documents such as those written by multidisciplinary teams.

Discussion
Key points

What is known about the topic

- Structural quality of protocols and clinical guidelines developed thus far for primary care is very low.
- Structural deficiencies of a tool designed to improve quality, such as a set of clinical guidelines, can invalidate the results obtained by implementing the guidelines.
- The main cause of inadequate quality is lack of knowledge by the developers and neglect of appropriate referencing.

What this study contributes

- Document quality varied significantly depending on the location of the center where the guidelines were developed.
- Quality was best in clinical guidelines developed by a multidisciplinary team and dealing with health problems normally covered by the center´s regular services.
- Poor structural quality of the documents we evaluated was manifested in different ways. Improvements are needed especially in the design and development of clinical guidelines for acute health problems.

A number of other opportunities for improvement were also identified, e.g., documents produced at certain health centers, those developed at nonteaching centers, guidelines that dealt with acute health problems, and guidelines included as an annex to a larger health program.

Directions for future research

Our study was undertaken as part of a research program with several areas for further development. We hope to develop corrective measures to improve the structural quality of clinical guidelines in primary health care in the region of Murcia, to re-evaluate the structure of this type of document throughout the region after corrective measures have been in place, and to validate a newly-developed instrument for the evaluation of clinical guidelines with regard to design, development, function and scientific evidence. In addition, future work will be aimed at evaluating the relevance of the guidelines, evaluating variability in the recommendations and in the scientific evidence on which they are based, and quantifying the actual use of these tools by professionals. We also hope to identify the characteristics clinical guidelines should have in order to be both useful and utilized, by comparing the structural quality criteria we used with those designed by

TABLE 4 **Characteristics that influenced structural quality of the protocols in the logistic regression analysis (n = 462)**

Variable	Category	*Odds ratio*	Significance	95% confidence interval
Health area (reference: Murcia)	Cartagena	0,19	< 0,001	0,07-0,51
	Lorca–Caravaca–Jumilla–Yecla	0,98	ns	0,37-2,60
	Cieza–Molina	0,30	< 0,001	0,14-0,65
Year center was opened	1988–1990	1,1	ns	0,50-2,43
	1990–1993	0,61	ns	0,24-1,57
Teaching accreditation (undergraduate and graduate)	Graduate	0,71	ns	0,26-1,95
	Nonteaching	0,53	ns	0,20-1,37
Professionals (multidisciplinary)	Physicians	0,14	< 0,001	0,06-0,34
	Nurses	0,67	ns	0,34-1,31
Population group (adult)	Children	1,61	ns	0,68-3,78
	Women	2,09	ns	0,86-5,06
	Elderly	0,74	ns	0,28-1,97
	Other	1,3	ns	0,39-4,24
Type of health problem (chronic)	Acute	0,20	< 0,001	0,09-0,46
	Others	0,40	< 0,05	0,17-0,90
Type of activity (acute care)	Education and prevention	0,6	ns	0,23-1,57
	Follow-up care	0,19	< 0,001	0,06-0,61
	Other	0,40	ns	0,14-1,10
Part of a program	Yes	0,24	< 0,001	0,12-0,47
Accredited	Not accredited	0,79	ns	0,41-1,49
Regular services	Yes	2,98	< 0,001	1,52-5,83

professional members of the primary care teams who will use these tools.

In conclusion, structural quality was associated with the proportion of criteria the document fulfilled and with compliance with more than the mean number of criteria. Quality varied across health care areas within the region of Murcia. Better guidelines were developed by multidisciplinary groups of authors, dealt with chronic health problems, referred to care provided as part of the center´s regular services, and were developed specifically as a tool for quality design and planning.

Better quality appeared to be associated, in some multivariate analyses, with teaching centers, with acute care, and with women´s health programs. Other characteristics were found to have room for improvement, e.g., guidelines developed at certain health care districts, at nonteaching centers, and those developed for acute health problems.

Acknowledgments

We thank all those persons who assisted us with this study, especially the members of the research group and members of the participating primary care teams.

References

1. Heather Palmer R. Evaluación de la asistencia ambulatoria. Principios y práctica. Madrid: Ministerio de Sanidad y Consumo, 1989.
2. Saturno PJ, Imperatori E, Corbella A. Evaluación de la calidad asistencial en atención primaria. Experiencias en el marco de la cooperación ibérica: diseños de proyectos. Madrid: Ministerio de Sanidad y Consumo, 1990.
3. Saturno PJ. Estrategias de implantación de las actividades de evaluación y mejora de la calidad. El ciclo evaluativo o ciclo de mejora de la calidad. En: Saturno PJ, editor. Curso de calidad asistencial en atención primaria. Madrid: Universidad de Murcia, DuPont Pharma, 1997.
4. Field MJ, Lohr KN. Clinical practice guidelines. Directions for a new program. Washington, D.C.: National Academy Press, 1990.
5. Schenbaum SC, Gottlieb LK. Algorithm based improvement of clinical quality. BMJ 1990; 301: 1374-1376.
6. Agency for Health Care Policy and Research. Guideline definition and attributes. Rockville: Department of Health and Human Services, 1990.
7. Bolster CJ. Quality research initiatives. Top Health Care Financ 1991; 18: 7-11.
8. Lomas J, Anderson GM, Dominick-Pierre K et al. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. N Eng J Med 1989; 321: 1306-1311.

9. Peter I, Tate J, Catchpole P. Practice activity analysis: collaboration between general practitioners and a family practitioner committee. J Royal Coll Gen Pract 1989; 39: 297-299.
10. Grol R. National standard setting for quality of care in general practice: attitudes of general practitioners and response to a set of standars. Br J Gen Pract 1990; 40: 367-351.
11. Moore SR. Pharmacy implications of clinical guideline development. Am J Hosp Pharm 1992; 49: 1763-1364.
12. American College of Emergency Physicians. Trauma care systems quality improvement guidelines. Ann Emerg Med 1992; 21: 736-739.
13. O'Brien JL, Hughes FX. The physician's role in quality assessment and improvement. Top Health Care Financ 1991; 18: 33-35.
14. Berwick DM, Henthoven A, Bunker JP. Quality management in NHS: The doctor´s role-II. BMJ 1992; 304: 304-308.
15. Saturno P, Saura J. Protocolos clínicos: una evaluación urgente y necesaria. Med Clin (Barc) 1994; 102: 717-718.
16. Saura Llamas J, Saturno Hernández PJ. Evaluación y mejora del diseño de los protocolos clínicos. Aten Primaria 1994; 13: 355-361.
17. Gaona Ramón JM, Saura-Llamas J, Saturno Hernández PJ, Romero Román JR. Evaluación del diseño de los protocolos clínicos de pediatría de la Comunidad Autónoma de Murcia. Pediatrika 1995; 15: 390-395.
18. Leal Hernández M, Saura-Llamas J, Saturno Hernández P. Evaluación de protocolos del Programa FMC. ¿Cumplen los criterios para la elaboración de protocolos clínicos? Aten Primaria 1997; 19: 426-430.
19. Saura-Llamas J, Saturno Hernández PJ, Gascón Cánovas JJ, López Montesinos MJ, Gaona Ramón JM y Grupo de Evaluación y Mejora de los Protocolos Clínicos. Evaluación de la calidad de los denominados «protocolos clínicos» de atención primaria elaborados en la Comunidad Autónoma de Murcia. Aten Primaria 1999; 23: 204-210.
20. Hazard B. Statistical methods for health care research. Filadelfia: Lippincot, 1997.
21. Kleinbaum DG, Kupper LL, Muller KE, Nizam A. Applied regression analysis and other multivariable methods (3.a ed.). Pacific Grove: Duxbury Press, 1998.
22. Hibble A, Kanka D, Pencheon D, Pooles F. Guidelines in general practice: the new Tower of Babel. BMJ 1998; 317: 862-863.
23. Cluzeau FA, Littlejohns P, Gimshaw JM, Feder G, Moran SE. Development and aplication of a generic methodology to assess the quality of clinical guidelines. Inter J Quality in Health Care 1999; 11: 21-28.
24. 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.

COMMENTARY

Problems with the evaluation of clinical guidelines and protocols

A.J. Jovell

Josep Laporte Library Foundation, Autonomous University of Barcelona, Spain.

The provision of health care that is based on clinical practice guidelines is the result of a process that should include these five stages: design, development, dissemination, implementation and analysis of impact.[1-3] Saura et al[4] have studied the quality of clinical practice guidelines with a set of criteria that assess elements of the design phase. The results of their study indicate that the best clinical practice guidelines were those designed by multidisciplinary groups, at teaching centers, within the context of a specific health program, dealing with a chronic health problem, and linked with services regularly provided by the center.

Now that the findings are in hand we should ask ourselves whether these criteria are appropriate to evaluate the quality of guidelines. It seems to me that they are necessary but not sufficient.

Firstly, an evaluation of clinical guidelines needs to distinguish between this type of instrument, intended as an aid clinical decision-making, and similar documents such as protocols, algorithms and referral pathways. This distinction avoids the so-called Hepburn bias, in which some people confuse Katharine Hepburn with Audrey Hepburn

Clinical practice guidelines

- The quality of clinical practice guidelines requires indicators that measure different parts of the process: design, development, dissemination, implementation and analysis of impact.
- In the evaluation of clinical practice guidelines form must not be confused with content, nor process with outcomes.
- Evaluating guidelines makes it possible to adapt them to specific health care contexts without the need to redesign them.
- Future research dealing with clinical practice guidelines should consider how to take patient preferences, health care costs and adaptation to individual cases into account. .

because the two have a certain feature (the surname, in this case) in common. Thus the process used to select guidelines for study should include those documents that fulfil the criteria of multidisciplinarity, scientific accreditation and, possibly, inclusion as part of a structured program of guideline design.[1] These criteria would make it possible to distinguish between clinical practice guidelines and other types of recommendations of lower quality.

Secondly, the evaluation of guidelines needs a wider set of criteria that take into account the different stages noted above. Generic evaluation instruments for this purpose have been validated, including the criteria proposed by the AGREE collaboration (www.agreecollaboration.org). This wider-ranging evaluation makes it possible to avoid confusing form with content, and process with outcomes. These two sources of confusion are the result of considering a guideline appropriate for good clinical practice (content) or for producing improvements in the quality of care (outcome) simply because it is well designed (form-process).

To avoid confusion between form and content, criteria based on scientific evidence should be used to judge the appropriateness of the recommendations in the guidelines for the clinical condition to be treated. In this connection, a clinical practice guideline can include multiple recommendations, some based on high-quality scientific evidence, and others based on low-quality evidence.[2] Preventing confusion between process and outcomes should involve an evaluation of the impact of clinical practice guidelines on improvements in health care. Thus, the value of the guidelines for improving clinical practice in terms of effectiveness and efficiency would ideally be assessed with instruments specifically designed for this purpose. In other words, guidelines should be examined with the aim of determining whether they are cost-effective. These issues cannot be resolved with evaluation criteria for structural quality, as such factors restrict the evaluation to aspects related with form and the design process.

Whether guidelines have the hoped-for impact on improvements in health care depends to a large extent on how the last three stages of the process—dissemination, implementation and impact analysis—are carried out. It is not enough to have good clinical practice guidelines; they must reach the professionals who will use them, and these practitioners must understand them. This is where the risk of the third type of confusion listed above (the so-called ods confusion) --attributing clinical relevance to results that may have statistical significance but lack clinical significance--comes into play.

Dissemination of the guidelines is favored by the progressive implementation of information and communication technologies in health care. Of special note are the existing databases specifically for clinical guidelines (www.fbjoseplaporte.org), which allow the texts to be consulted from anywhere. In addition to dissemination of the guidelines, the need to design strategies to promote their implementation is evident. Implementation constitutes the true Achilles´ heel of the process, as simply having guidelines is not enough–they must also be used so that they can fulfil their mission of helping to improve the quality of care. In this connection, the fact the good quality guidelines tended to deal with problems that are covered by services normally offered by a given center[4] may be related with the presence of incentives that favor the use of guidelines. A good implementation process involves appropriate training for professionals in the interpretation, evaluation, adaptation and utilization of clinical practice guidelines, as well as the development of incentives that favor their use. The absence of such incentives may favor the use of defensive medicine strategies, and may favor variability in clinical practice. Thus suitable strategies must be designed to create incentives for the appropriate use of guidelines as instruments to improve clinical practice.[5,6]

The process of developing clinical practice guidelines, which includes all stages from design to implementation, is expensive. It is therefore inefficient to create multiple guidelines for the management of the same clinical condition –a phenomenon known as guidelinemania–. The future development of guidelines should not aim to create additional guidelines, but to evaluate existing ones and adapt them, with suitable protocols, to specific health care contexts.[3] Moreover, research on clinical practice guidelines should take into account issues related with patients' preferences, and the costs to health care budgets involved in their design. In this connection it should be recalled that clinical practice guidelines are simply instruments designed to aid decision-making, and that scientific evidence is one of the many factors that influence clinical decision-making. Thus, in addition to scientific evidence, the evidence from the clinical case at hand needs to be considered, which means that the clinical practice guideline must be adapted to a specific clinical history. The physician´s role is essential to ensure that the guidelines are applied correctly to each individual patient´s case. This is why good physicians, as well as good guidelines, are needed.

The results of the study that appears in this issue of Atención Primaria[4] illustrate the differences in quality of clinical practice guidelines as evaluated with a small set of structural criteria. Of note is the lack of data on how (or whether) the guidelines were updated, this being one of the prime features of such documents. As scientific evidence becomes outdated, updating the guidelines will depend on well-planned reviews of the available knowledge. As noted at the beginning of this article, it should be recalled that it is just as difficult to find a well-designed protocol as it is to find a well-designed clinical guideline.

References

1. Jovell AJ, Navarro-Rubio MD. Guías de práctica clínica. FMC 1995; 2: 152-156.

2. Jovell AJ, Navarro-Rubio MD, Aymerich M, Serra-Prat M. Metodología de diseño y elaboración de guías de práctica clínica en atención primaria. Aten Primaria 1997; 20: 259-266.
3. Jovell AJ, Navarro-Rubio MD, Aymerich M. Guías de práctica clínica. En: Jovell AJ, Aymerich M, editores. Evidencia científica y toma de decisiones en sanidad. Barcelona: Acadèmia de Ciències Mèdiques de Catalunya i Balears, 2000.
4. Saura-Llamas J, Saturno Hernández PJ, Romero Román JR, Gaona Ramón JM, Gascón Cánovas JJ y Grupo de Evaluación y Mejora de los Protocolos Clínicos. Características de las guías clínicas de atención primaria que se asocian a una mayor calidad asistencial. Aten Primaria 2001; 28: 525-534.
5. Grimshaw J, Freemantle N, Wallace S, Rusell I, Hurwitz B, Watt I et al. Developing and implementing clinical practice guidelines. Quality in Health Care 1995; 4: 55-64.
6. Woolf SH, Grol R, Hutchinson A, Eccles, Grimshaw J. Clinical guidelines. Potential benefits, limitations, and harms of clinical guidelines. BMJ 199; 318: 527-530.