



ORIGINAL ARTICLE

Using Appraisal of Guidelines Research and Evaluation to Appraise Nursing Clinical Practice Guidelines in Taiwan and to Compare Them to International Studies

Kee-Hsin Chen^{1,2,3,4}, Ching-Chiu Kao^{1,3,4*}, Hsueh-Erh Liu^{5*}, Wen-Ta Chiu⁶, Ken N. Kuo^{4,7}, Chieh-feng (Cliff) Chen^{4,8,9,10*}

¹ Department of Nursing, Taipei Medical University, Wan Fang Hospital, Taiwan

² Graduate Institute of Clinical Medical Sciences, College of Medicine, Chang Gung University, Taiwan

³ School of Nursing, College of Nursing, Taipei Medical University, Taiwan

⁴ Center for Evidence-Based Medicine, Taipei Medical University, Taiwan

⁵ School of Nursing, Chang Gung University, Taiwan

⁶ Department of Health, Executive Yuan, Taiwan

⁷ Division of Health Policy Research and Development, Institute of Population Health Sciences, National Health Research Institutes, Taiwan

⁸ Department of Public Health, School of Medicine, College of Medicine, Taipei Medical University, Taiwan

⁹ Department of Surgery, Taipei Medical University, Wan Fang Hospital, Taiwan

¹⁰ Evidence-Based Medicine Center, Taipei Medical University, Wan Fang Hospital, Taiwan

ARTICLE INFO

Article history:

Received: Dec 27, 2010

Accepted: Sep 1, 2011

KEY WORDS:

evidence-based medicine;
nursing clinical practice guidelines

Background: Implementation of clinical practice guidelines (CPGs) can reduce medical practice variations and enhance effectiveness. In 2008, the Taiwan Department of Health delegated the development of CPGs on six health topics. By February 2010, these CPGs had been completed.

Methods: The Appraisal of Guidelines Research and Evaluation (AGREE) instrument was used to assess the quality of the guidelines. We organized groups to develop CPGs and called on 131 experts with various specialties.

Results: The average AGREE scores for the six major domains were 86% (scope and purpose), 63% (stakeholder involvement), 77% (rigor of development), 72% (clarity and presentation), 53% (applicability), and 69% (editorial independence).

Conclusion: We recommend the following: (1) opinions from all stakeholders, especially patients, should be considered to improve the content of the guidelines; and (2) health education resources, audit forms, and other tools should be emphasized more in the development of CPGs.

Copyright © 2011, Taipei Medical University. Published by Elsevier Taiwan LLC. All rights reserved.

1. Introduction

Clinical practice guidelines (CPGs) are systematically defined statements to assist practitioners and patients in making decisions regarding appropriate health care in specific circumstances. Implementation of such guidelines can reduce medical practice variations and enhance effectiveness and efficiency while at the

same time achieving cost-effectiveness.¹ Following its definition of CPGs in 1990, the Institute of Medicine (IOM) introduced the first guideline appraisal tool in 1992.²

The US Department of Health and Human Services established the National Guideline Clearinghouse (NGC) in 1998. The UK, Australia, and New Zealand, along with other countries, also developed many CPGs to assist healthcare professionals in applying research evidence to clinical practice. It has been difficult to apply these guidelines in other countries because of cultural differences and the lack of motivation, awareness, familiarity, agreement and training in the use of the guidelines, and language barriers.^{3–5}

The National Health Research Institutes (NHRI) in Taiwan established the Taiwan Cooperative Oncology Group (TCOG), a research institute that has been developing an early practice guideline program since 1997 and has published CPGs for cervical and breast cancer. In 2005, the Department of Health (DOH) in

* Corresponding authors. (1) CCK, Department of Nursing, Taipei Medical University, Wan Fang Hospital, 111 Xinglong Road, Section 3, Wenshan District, Taipei 116, Taiwan. (2) HEL, School of Nursing, Chang Gung University, 259 Wen-Hwa 1st Road, Kwei-Shan, Taoyuan 333, Taiwan. (3) CC, Evidence-Based Medicine Center, Taipei Medical University, Wan Fang Hospital, 111 Xinglong Road, Section 3, Wenshan District, Taipei 116, Taiwan.

E-mail addresses: C.-C. Kao <675031@gmail.com>; H.-E. Liu <sarah@mail.cgu.edu.tw>; C. Chen <clifchen@tmu.edu.tw>

Taiwan began developing CPGs; however research data on the quality of these guidelines have been minimal.

Since February 2008, the DOH has delegated the development of CPGs on high-priority topics in nursing to healthcare institutions, requiring significant outcomes, effectiveness, and benefits to be on a par with those of international CPG development centers.⁶ By February 2010, six CPGs had been completed.

In this study, we analyzed the quality of the six CPGs developed by nurses in Taiwan in order to validate their contents before recommending their application to clinical practice. Additionally, the data were compared to international guidelines in order to understand common interests, differences, and trends, if any. The results of this study can serve as preliminary quality information on the configuration of CPGs and guide subsequent development as well as policy making, while providing a reference for related future studies.

2. Methods

2.1. Nursing CPG development

The development process consisted of several stages. First, we organized guideline-developing groups (GDGs) for these CPGs. In total, 131 experts with specialties in various fields were enlisted to develop the CPGs. Second, the GDGs were introduced to the topic, scope, format, content, and the Appraisal of Guidelines Research and Evaluation (AGREE) instrument. Third, GDGs performed systematic literature reviews and drafted the CPGs. Fourth, meetings with experts were organized to review and finalize the CPGs. Each completed CPG was reviewed by three experts who appraised the content in accordance with AGREE. The rating of the recommendations was based on the 2004 Taiwan Guideline Developer's Handbook.⁷ We also surveyed the clinical applicability with personnel in 15 hospitals nationwide, including 568 registered nurses and respiratory therapists. The results showed that 89.9% of nurses thought these CPGs could help them provide more effective and higher-quality care as well as health education.

2.2. Methods of guideline appraisal

2.2.1. Structure and content of AGREE

In this study, the AGREE instrument was used as a guideline quality-assessment tool. The AGREE instrument is an internationally validated tool with a rigorously developed methodology.⁸ It contains six domains with a total of 23 items and allows for the assessment of several components that are integral to guideline development. It can serve as a standard to guide development, implementation, and supervision, and can also be used to compare different international guidelines.⁹ AGREE was translated into traditional Chinese according to the methodology in Guillemin et al.'s "Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines" in 2006 and was validated.¹⁰

When AGREE is used to assess guidelines, there should be at least two reviewers for each guideline.¹¹ Each item is rated on a 4-point scale: 4, "strongly agree"; 3, "agree"; 2, "disagree"; and 1, "strongly disagree". Each item has a comment column for the reviewer to explain why a certain response was given. The mathematical formula of the agreement (%) is shown in Figure 1.

Adding all the scores of the individual items in a domain and then standardizing them across reviewers as follows was used to obtain the score for each domain: (obtained score – minimal possible score)/(maximal possible score – minimal possible score). The maximal possible score for each domain was the maximal number of items multiplied by the number of reviewers times 4

$$\text{Agreement (\%)} \text{ of an item among reviewers} = \frac{\text{Count of Scale (3), (4)}}{\text{Count of Scale (1), (2), (3), (4)}}$$

Each item is rated on a 4-point scale:
(4) = "strongly agree"; (3) = "agree"; (2) = "disagree"; and (1) = "strongly disagree."

Figure 1 Calculating agreement (%).

(i.e., the score for "strongly agree"). The minimal possible score for each domain was the minimal number of items multiplied by the number of reviewers times 1 (i.e., the score for "strongly disagree"). Domain scores were calculated by summing up all scores of the individual items in a domain and standardizing the total as a percentage of the maximum possible score for that domain. The mathematical formula is given in Figure 2.

Regarding the overall assessment of each guideline, if the majority scores for the six domains were >60%, it was "strongly recommended". If the majority scores were 30–60%, it was "recommended with provisos". If the majority scores were <30%, the appraisal result was "would not recommend". At the end of the assessment process, reviewers were asked to give an overall comprehensive appraisal, choosing from "strongly recommended", "recommended" (with provisos or alterations), "would not recommend", and "unsure", according to the AGREE statement.¹²

After the CPGs were drafted, each CPG was sent to three reviewers for assessment according to AGREE. Reviewers were composed of a group of well-experienced guideline experts who were recommended by the Taiwan Evidence Based Medicine Association (TEBMA). The review training process was based on instructions of the *AGREE Instrument Training Manual*.¹² All reviewers conducted the review process independently. Reviewers completed the rating process and gave overall suggestions after carefully reading the rating instructions provided in the AGREE manual.

3. Results

The standardized mean scores for the six Taiwanese CPGs in the six major domains of AGREE were 86%, 63%, 77%, 72%, 53%, and 69%. The detailed scores for each item and the overall comprehensive appraisal of the six CPGs are given in Table 1.

The mean score of the domain "scope and purpose" was 86% (standard deviation, SD 10.2%; range, 70–96%); the mean score of the domain "rigor of development" was 77% (SD 5.2%; range, 70–84%); and the mean score of the domain "clarity and presentation" was 72% (SD 8.4%; range, 61–81%). In these three domains, the standardized scores for the six guidelines were all >60%. In addition, the mean score of the domain "stakeholder involvement" was 63% (SD 10.6%; range, 47–72%), and four guidelines (66.7%) scored higher than 60%. The mean score of the domain "applicability" was 53% (SD 15.1%; range, 37–78%), which was the lowest

$$\text{Mean domain score} = \frac{\text{Obtained score} - \text{Minimal possible score}}{\text{Maximal possible score} - \text{Minimal possible score}}$$

Maximal possible score = 4 (strongly agree) * (items) * (number of reviewers)
Minimum possible score = 1 (strongly disagree) * (items) * (number of reviewers)

Figure 2 Calculating domain score.

Table 1 AGREE rating results for Taiwan nursing clinical practice guidelines (CPGs)

	Scope and purpose	Stakeholder involvement	Rigor of development	Clarity and presentation	Applicability	Editorial independence	Rev 1	Rev 2	Rev 3
CPGs on gastrointestinal feeding of critically ill patients	96	70	84	81	44	61	SR	R	R
CPGs for elderly people with constipation	93	53	73	72	52	50	SR	R	R
CPGs for children with a fever	85	72	79	61	63	94	R	R	R
CPGs for prevention of urinary tract infections in adults with long-term indwelling catheters	93	69	81	81	78	72	R	SR	R
CPGs on fall prevention among hospitalized adults	70	69	70	75	37	78	R	SR	R
CPGs for the maintenance of respiratory function	78	47	76	64	44	61	R	R	R
Mean scores of Taiwanese CPGs	86	63	77	72	53	69	/	/	/

Rev 1, 2, and 3, reviewers number 1, 2, and 3; SR, strongly recommended; R, recommended.

score of all domains. Only two guidelines (33.3%) had a standardized score of >60%. The mean score of the domain “editorial independence” was 69% (SD 15.5%; range, 50–94%), and five guidelines (83.3%) scored >60%. Except for the domain of applicability, the mean scores of the domains were >60%. Thus, the domain of applicability had room to improve.

4. Overall comparison of AGREE rating results for Taiwan with international results

The same AGREE instrument was used internationally in 2005 to rate 17 pediatric CPGs, and scores for the six major domains were 84%, 42%, 54%, 78%, 19%, and 40%. Fourteen of the guidelines were worth recommending.¹³ Compared to those results, Taiwanese guidelines appeared to have superior scores for the domains “stakeholder involvement”, “rigor of development”, “applicability”, and “editorial independence”, but an inferior AGREE score for the domain “clarity and presentation” (Figure 3).

5. Agreement among the three reviewers for AGREE instrument items

Results of the AGREE rating for CPGs in our study showed over 80% agreement among the experts, and there were four items with agreement of 41–60% and four with agreement of 61–80%. The overall trend was similar to results of international studies.¹³ However, two items were lower than all items of the international report and were 21–40%: items 5 (patients’ views and preferences have been sought) and 23 (conflicts of interest of guideline development members have been recorded) (Table 2).

6. Discussion

An overview of the scores in each domain for the CPGs from Boluyt et al.’s study¹³ showed similar results for the mean scores of the domains “scope”, “purposes”, and “clarity and presentation”. The score of the Taiwanese CPG for the domain “clarity and presentation” was slightly worse than the international scores, while the scores in the domains “stakeholder involvement”, “rigor of development”, “applicability”, and “editorial independence” were better. Our program had already introduced the AGREE instrument soon after recruiting members for the GDGs; therefore, the description of the drafting process was in accordance with the description in the AGREE instrument. In contrast, the study by Boluyt et al. included CPGs obtained from the internet, and AGREE was not necessarily followed during the development of those CPGs. This may have been the reason that the international CPGs had lower scores in most domains.

Scores of the two studies were high in both domains of “scope and purpose” and “rigor of development”. This indicates thorough consideration of describing the applicable scope and purpose of the guidelines, utilization of systematic methods to collect literature and evidence, and description of criteria for selecting evidence and formulating the recommendations were used.

For item 7 of this domain (the guideline has been pilot-tested among target users), the scores were low in both the Taiwanese and the international guidelines. This implies that it is difficult to complete pilot studies and efficacy assessments before announcing the guidelines.

In the domain “clarity and presentation”, the CPGs in our study did not provide a brief version of a care flow chart, health education materials, or tools (such as compact discs), so the score for item 18 (the guideline is supported by tools for its application) was lower.

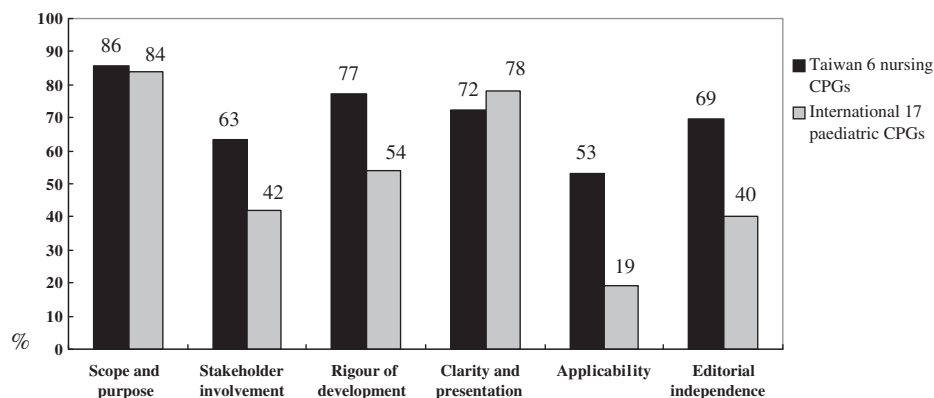


Figure 3 International comparison of AGREE rating results.

Table 2 Agreement among reviews for AGREE instrument items

Agreement (%)	No. of items among the Taiwanese six nursing CPGs	No. of items among the International 17 paediatric CPGs ¹³
0	0	0
0–20	0	0
21–40	2	0
41–60	4	3
61–80	4	12
81–100	13	8

CPGs = clinical practice guidelines.

Finally, the lowest standardized scores in both studies were in the domain “applicability”. The percentages of domestic and international CPGs that had not been subjected to pilot studies in an early stage of the announcement were high, and it is possible that a substantial description is difficult for items 19 (the potential organizational barriers in applying the recommendations have been discussed), 20 (the potential cost implications of applying the recommendations have been considered), and 21 (the guideline presents key review criteria for monitoring and/or auditing purposes). This may explain the lower scores for this domain. The main purpose of promoting CPGs is to enhance the quality of care; therefore, it is recommended that major assessment and monitoring criteria should be clearly defined and described when CPGs are established in the future in order to provide more-substantial guidelines for monitoring the quality of care.

7. From development of CPGs to knowledge translation

From the analysis of the quality of the contents of CPGs in this study, preliminary results showed that scores for the domains of “scope and purpose”, “rigor of development”, and “clarity and presentation” were relatively high. The contents of these domains have to do with extraction, induction, and translation of knowledge and are part of knowledge synthesis. This indicates that implementing guidelines is perceived as more difficult than developing them.

From the appearance of new evidence in the literature to actually putting it into practice in clinical work, clinical professionals must pass through various stages, including “being aware of it”, “accepting it”, “making it applicable”, “being able to do it”, “acting on it”, “agreeing with it”, and “adhering to it”.¹⁴ In clinical settings, there may be interference from factors, such as costs and patient acceptance and expectations of new measures that affect whether or not new knowledge becomes mainstream opinion.¹⁵ As a result, significant knowledge leaks occur during the process, and what is actually applied to patient care is often far less than what is indicated in the research literature.

In domestic and international AGREE scoring, the domains “stakeholder involvement” and “editorial independence” address how guideline development experts apply knowledge in practice. CPGs are useful tools in the process of knowledge translation. Putting more efforts into these two domains as well as the domain of “applicability” (which had the lowest score) may reduce barriers to clinical application and bridge the gap between empirical knowledge and clinical practice to enhance the quality of care.

8. Limitations of this study

1. Ideally, it would be better to compare Taiwan nursing guidelines to international nursing guidelines if available. However, due to the limited amount of related nursing research, we chose to compare the AGREE scores of international guidelines to Taiwanese nursing guidelines in this study.

2. Evidence-based practice features the “3Es”: evidence, experts, and expectations. An analysis of the quality of CPGs initially showed that clinicians are quite skillful in processing empirical knowledge, but evidence is only a part of clinical care. Minimizing the difference between empirical knowledge and practice so that healthcare professionals are willing to change their care patterns and provide evidence-based health care, while taking patients’ opinions into consideration, is the common goal of guideline developers and clinicians. The development process of Taiwanese nursing CPGs did not invite patient representatives, which might be another limitation.

9. Conclusion

The six Taiwanese CPGs can be recommended for clinical practice based on the AGREE assessment. There are two components of knowledge translation: “knowledge synthesis” and “clinical application”. In terms of overall trends, the technology of knowledge synthesis is gradually maturing, but more efforts are required for knowledge application, and the development of CPGs can serve as a bridge between the two. For future development of CPGs, we recommend the following: (1) upon formation of GDGs, opinions from all stakeholders, especially patients and primary caregivers, should be considered to make the contents of the guidelines more complete and feasible; (2) because clinical work can be busy and complex, guidelines that come with simple care flow charts, health education resources, audit forms, and other tools should be given more emphasis in future guideline development; and (3) the quality of CPG development could be enhanced by ensuring that items examined by the AGREE instrument are appropriately addressed.

References

- Field MJ, Lohr KN. *Clinical practice guidelines: directions for a new program*. Institute of Medicine. Washington, D.C: National Academies Press; 1990.
- Field MJ, Lohr KN. *Clinical practice guidelines: directions for a new program*. Institute of Medicine. Washington, D.C: National Academies Press; 1992.
- Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, Rubin HR. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA* 1999;**282**:1458–65.
- Cabana MD, Ebel BE, Cooper-Patrick L, Powe NR, Rubin HR, Rand CS. Barriers pediatricians face when using asthma practice guidelines. *Arch Pediatr Adolesc Med* 2000;**154**:685–93.
- Gravel K, Legare F, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: a systematic review of health professionals' perceptions. *Implement Sci* 2006;**1**:16.
- Council of Europe. Developing a methodology for drawing up guidelines on best medical practices. *Z Arztl Fortbild Qualitatssich* 2002;**96**(suppl. 3):5–59 (Recommendation (2001) 13 and explanatory memorandum).
- AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Qual Safe Health Care* 2003;**12**:18–23.
- Vlayen J, Aertgeerts B, Hannes K, Sermeus W, Ramaekers D. A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit. *Int J Qual Health Care* 2005;**17**:235–42.
- Bureau of National Health Insurance, Department of Health, Executive Yuan, Taiwan Joint Commission on Hospital Accreditation and National Health Research Institutes. *Taiwan guideline developer's handbook*. Available from: http://www.nhi.gov.tw/webdata/AttachFiles/Attach_613_1_CPG.pdf; 2004 [accessed 20.05.2010].
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993;**46**:1417–32.
- The AGREE Collaboration. *The Original AGREE Instrument Translations*. Available from: <http://www.agreetrust.org/resource-centre/the-original-agree-instrument/the-original-agree-instrument-translations/>; 2001.
- The AGREE Collaboration. *AGREE instrument training manual*. Available from: <http://www.agreecollaboration.org/pdf/aitraining.pdf>; 2003.
- Boluyt N, Lincke CR, Offringa M. Quality of evidence-based pediatric guidelines. *Pediatrics* 2005;**115**:1378–91.
- Lang ES, Wyer PC, Haynes RB. Knowledge translation: closing the evidence-to-practice gap. *Ann Emerg Med* 2007;**49**:355–63.
- Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;**362**:1225–30.