Postpublication External Review of the Japanese Guidelines for the Management of Stroke 2004

Yukito Shinohara, MD, PhD; Masao Nagayama, MD, PhD; Hideki Origasa, PhD

Background and Purpose—Many guidelines for management of stroke have been published throughout the world, but no postpublication external review of any set of stroke guidelines by users, using standard checklists, has been reported. The purpose of this article is to present the results of an external review of the Japanese Guidelines for the Management of Stroke 2004, conducted several months postpublication.

Methods—Forty-one evaluators, who had not been involved in developing the guidelines, were selected from representative stroke centers and institutions in Japan. They consisted of 30 physicians including 22 stroke specialists, and 11 nurse practitioners. Three standard checklists, ie, Appraisal of Guidelines for Research and Evaluation (AGREE) instrument, checklist by Shaneyfelt et al, and the Conference on Guideline Standardization (COGS) checklist, were used.

Results—Confidence ratios according to the AGREE checklist were 75%, 77%, and 86% for stroke specialists, physicians other than stroke specialists, and nurse practitioners, respectively. The average scores were 2.98, 3.13, and 3.29, respectively. The confidence ratios according to the checklist by Shaneyfelt et al were 72%, 73% and 86% respectively, and those for the COGS checklist were 66%, 74%, and 91%, respectively.

Conclusions—Although it is impossible to compare our results with those for other stroke guidelines, because none of them has been externally reviewed by users postpublication, our results seem better than those for published guidelines for treatment of other diseases in Japan. These results should be helpful in the process of updating stroke guidelines in Japan and elsewhere. (Stroke. 2009;40:1439-1443.)

Key Words: stroke ▪ stroke guidelines ▪ evidence-based medicine ▪ AGREE ▪ COGS ▪ external review of guidelines

Guidelines for the management of stroke were first published in 1994 by the American Heart Association,1 and these were followed by the evidence-based guidelines of the Royal College of Physicians in the United Kingdom,2 and others. However, there are differences in approved therapeutic agents and stroke subtypes between Japan and other countries, as well as racial differences, so we set out to create the first Japanese stroke guidelines, taking into consideration the detailed evidence and experience accumulated in Japan.

Representatives from 5 stroke-related societies in Japan (the Japan Stroke Society, the Japan Neurosurgical Society, Japanese Society of Neurology, Japanese Society of Neurological Therapeutics, and the Japanese Association of Rehabilitation Medicine) gathered in Tokyo in October 2000 and set up the Joint Committee on Guidelines for the Management of Stroke (Chairperson: Yukito Shinohara, Vice-Chairpersons: Takashi Yoshimoto, Yasuo Fukuchi, Shigenobu Ishigami, total of 107 stroke specialists). The Committee was divided into 5 chapters focused on stroke in general at the extreme acute stage (before the definite diagnosis of stroke subtypes, including prevention), cerebral infarction and TIA, intracerebral hemorrhage, subarachnoid hemorrhage, and rehabilitation. The guidelines were produced on the basis of detailed and repeated discussions among the chairperson, vice-chairpersons, members, reviewers, a clinical epidemiologist (Hideki Origasa), and the Secretary-general (Masao Nagayama). More than 110 000 references (mainly written in English, but including some reports in the Japanese literature from 1992 to 2003, searched by using MEDLINE, Cochrane Library, Japan Centra Revuo Medicina, and other resources) were reviewed. Work on the Japanese Guidelines for the Management of Stroke 2004 was completed around the end of 2003, and the first edition was published in February 2004.3,4 This edition of the guidelines is only for physicians, including stroke specialists, and stroke nurses. Although the guidelines are written in Japanese, an outline has been reported in English5 and the difference between the Japanese guidelines and Western guidelines has also been discussed in part.6 Many guidelines are reviewed by external reviewers during their preparation,3,4 but it...
Table 1. Classification of Evidence Levels by the Joint Committee on Guidelines for the Management of Stroke

<table>
<thead>
<tr>
<th>Levels of Evidence</th>
<th>Type of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ia</td>
<td>Meta-analysis (with homogeneity) of randomized controlled trials</td>
</tr>
<tr>
<td>Ib</td>
<td>At least one randomized controlled trial dealing with a large series of subjects</td>
</tr>
<tr>
<td>IIa</td>
<td>At least one well-designed controlled study but without randomization</td>
</tr>
<tr>
<td>IIb</td>
<td>At least one well-designed quasi-experimental study</td>
</tr>
<tr>
<td>III</td>
<td>At least one well-designed nonexperimental descriptive study (eg, comparative studies, correlation studies, case studies)</td>
</tr>
<tr>
<td>IV</td>
<td>Expert committee reports, opinions, or experience of respected authorities</td>
</tr>
</tbody>
</table>

This classification is based on the classification of National Clinical Guidelines for Stroke (1999) which was used by the Royal College of Physicians, and also partly modified based on the classification of Oxford Center for Evidence-based Medicine (2001).

is also important that guidelines should be reviewed externally after publication by people who have used them but who did not participate in their preparation. However, to our knowledge, no such postpublication external review of stroke guidelines using standard checklists has yet been conducted. Here, we present the results of a postpublication external review of the Japanese Guidelines for the Management of Stroke 2004, performed several months postpublication by 41 independent evaluators and designed to investigate particularly the appropriateness of our guidelines, as judged by various user groups.

Materials and Methods

Process of Production of Evidence-Based Stroke Guidelines

The actual process of the production of our evidence-based stroke guidelines can be summarized briefly as follows: (1) foundation of the Committee; (2) evaluation of key themes and extraction of relevant issues; (3) search of the literature; (4) critical appraisal of the relevant reports; (5) determination of the evidence level for all the reports to be cited, and preparation of the evidence table; (6) recommendation grading for each of the themes or issues; (7) reevaluation of the validity of the statements in the guidelines by at least 3 external reviewers during the preparation process; and (8) publication in book form and on the web.

The evidence levels of all the reports cited were determined based on the criteria produced by our Committee (Table 1). The grade of recommendation for each theme or problem was also determined based on our own criteria (Table 2).

Table 2. Classification of Recommendation Grade by the Joint Committee on Guidelines for the Management of Stroke

<table>
<thead>
<tr>
<th>Grades of Recommendations</th>
<th>Type of Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Strongly recommended (At least one report at evidence level Ia or Ib)</td>
</tr>
<tr>
<td>B</td>
<td>Recommended (At least one report at evidence level IIa or IIb)</td>
</tr>
<tr>
<td>C1</td>
<td>Can be considered, although scientific evidence is inconclusive (evidence level III or IV)</td>
</tr>
<tr>
<td>C2</td>
<td>Not recommended due to lack of scientific evidence</td>
</tr>
<tr>
<td>D</td>
<td>Not recommended due to evidence of ineffectiveness</td>
</tr>
</tbody>
</table>

stroke in the acute and chronic stages, including prevention; (6) it is not a translation of stroke guidelines from the West, but rather a new guideline for the treatment of Japanese people, developed by Japanese doctors, and taking into consideration the evidence from the Japanese language literature, as well as the local situation in Japan; (7) novel pharmaceutical treatments so far approved only in Japan, but for which only one randomized controlled trial dealing with a rather small number of subjects was available, were tentatively assigned recommendation grade B; (8) original classifications based on global standards were used for evidence levels and recommendation grades; (9) an evidence level was assigned to all the cited literature; and (10) an attempt was made to clarify what evidence is lacking in each field of stroke.

Postpublication External Review Using Standard Checklists

To assess the validity and quality of the Japanese Guidelines for the Management of Stroke 2004, we commissioned an external review several months after publication, when the guidelines were in widespread use in Japan. The assessment was made as one of the Research on Health Technology Assessments of the Ministry of Health, Labor, and Welfare, Japan. Forty-one evaluators, who had not been involved in developing the guidelines, were randomly selected from representative stroke centers and institutions in Japan. They consisted of 30 physicians, of whom 22 were stroke specialists (12 neurologists, 6 rehabilitation specialists, and 4 neurosurgeons), and 11 nurse practitioners, and we asked them to independently fill in a series of questionnaires based on three current international standard checklists for the appraisal of clinical practice guidelines: (1) checklist of Appraisal of Guidelines for Research & Evaluation (AGREE) instrument; (2) checklist by Shaneyfelt et al; (3) the Conference on Guideline Standardization (COGS) checklist.

The AGREE instrument was developed by a group of researchers from 13 countries to provide a systematic framework for assessing guideline quality. This instrument was thoroughly evaluated and refined and is now a commonly used assessment instrument. AGREE consists of 23 key items organized in 6 domains, that is, “Scope and purpose,” “Stakeholder involvement,” “Rigour of development,” “Clarity and presentation,” “Applicability,” and “Editorial independence.” Each item is rated on a 4-point scale ranging from 4 “Strongly Agree” to 1 “Strongly Disagree,” with 2 intermediate points: 3 “Agree” and 2 “Disagree.” The scale measures the extent to which a criterion (item) has been fulfilled.

Shaneyfelt et al captured key features of guideline development and created a 25-item checklist, using a “yes” or “no” format, to measure adherence to these elements, broadly grouped into “standards on guideline format and development,” “identification and summary of evidence,” and “formulation of recommendations.” These items were piloted at workshops and pretested by the authors and were endorsed as comprehensive and valid.
COGS created a 18-item checklist, using a “yes” or “no” format, to promote guideline quality and facilitate implementation, and also to be used prospectively by guideline developers.

**Results**

**Checklist of the AGREE Instrument**

To allow comparison of the results obtained with the 3 checklists, the confidence ratio for the guidelines was calculated from the answers as the proportion of those evaluators who answered either “Strongly agree” or “Agree.” Confidence ratios according to the AGREE checklist, and the average AGREE scores and standard errors are shown in Table 3.

Regarding the results of item-by-item analysis, among the leading 5 items that received favorable appraisals, the following 3 items were common to all the subgroups: “the patients to whom the guideline is meant to apply are specifically described (scores 3.71, 3.71, and 3.73 for stroke specialists, physicians other than stroke specialists, and nurse practitioners, respectively),” “the clinical question(s) covered by the guideline is (are) specifically described (3.60, 3.71, and 3.83, respectively).” Among the unfavorable appraisals of the guidelines, “the guideline has been piloted among target users” was common to all the subgroups (2.45, 2.38, and 2.27, respectively). Among the leading 5 items of the unfavorable appraisals, the items “the health benefits, side effects, and risks have been considered in formulating the recommendations” (3.00) and “the potential organizational barriers in applying the recommendations have been discussed” (2.91) were mentioned specifically by stroke specialists. The items “the recommendations are specific and unambiguous,” “the patients’ views and preferences have been sought,” and “the target users of the guideline are clearly defined” were mentioned specifically by nurse practitioners (3.18, 2.55, and 2.91, respectively).

The Figure illustrates the results of analysis by domain scores. Domain score was higher in “scope and purpose” and relatively low in “applicability,” “stakeholder involvement,” and “editorial independence.”

**Checklist by Shaneyfelt et al**

The confidence ratio for the guidelines was calculated from answers for all the 25 items as a proportion of those evaluators who answered “yes.” Confidence ratios according to this checklist are shown in Table 3.

Regarding the confidence ratio for each item, consistently favorable appraisals (>90% in all of the subgroups and a difference of <10% among them) were obtained for the following items: “rationale and importance of the guideline are explained,” “the participants in the guideline development process and their areas of expertise are specified,” “method of identifying scientific evidence is specified,” “the evidence used is identified by citation and referenced.” Unfavorable appraisals (<50% in all of the subgroups) were observed for the following items: “the method by which the

<table>
<thead>
<tr>
<th>Category</th>
<th>Average score</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCOPE AND PURPOSE</td>
<td>3.56</td>
<td>0.89</td>
</tr>
<tr>
<td>STAKEHOLDER INVOLVEMENT</td>
<td>2.78</td>
<td>0.70</td>
</tr>
<tr>
<td>RIGOUR OF DEVELOPMENT</td>
<td>3.33</td>
<td>0.83</td>
</tr>
<tr>
<td>CLARITY AND PRESENTATION</td>
<td>3.05</td>
<td>0.76</td>
</tr>
<tr>
<td>APPLICABILITY</td>
<td>2.75</td>
<td>0.69</td>
</tr>
<tr>
<td>EDITORIAL INDEPENDENCE</td>
<td>2.84</td>
<td>0.71</td>
</tr>
</tbody>
</table>
On the other hand, dissociated appraisals between subgroups (difference of confidence ratio >40% between any 2 subgroups; this cut-off point 40% was arbitrarily selected as indicating substantial disagreement) were noted for the following items: “the health outcomes are specified,” “benefits and harms of specific health practices are specified,” and “benefits and harms are quantified.” For all of these items, stroke specialists showed relatively low confidence ratios compared with others, especially nurse practitioners (Table 4).

**Discussion**

To assess the validity and quality of our guidelines, we commissioned a postpublication external review by using 3 current international checklists which were developed for the appraisal of clinical practice guidelines. Although it is impossible to compare our results with those for other stroke guidelines in the world, because prior stroke guidelines in other countries have not been subjected to this kind of external review, our results were better than those for published guidelines for treatment of other diseases in Japan, and the confidence ratio reached 66% to 75% among stroke specialists, 73% to 77% among physicians other than stroke specialists, and 86% to 91% among nurse practitioners (Table 3). It should be noted that, although none of the 41 evaluators had been involved in developing our guidelines at any way, 18 of them were from a facility to which a member of the Joint Committee on Guidelines also belonged.

Several problems with our guidelines have become apparent, concerning items about pilot study or external review before publication, description of costs and patient preferences, and so on. In some cases, appraisals differed among stroke specialists, physicians other than stroke specialists, and nurse practitioners. Appraisals by stroke specialists were generally less positive than those by the other groups for many items, including those about funding source/sponsor, pilot study or external review, potential benefits and harms, and so on.
algorithm, implementation considerations, and applicability. Nurse practitioners were less positive than the others regarding items about patient preferences, target users, and the clarity of the recommendations.

Several months after the publication of our guidelines, a Japanese publishing company conducted an independent survey about our guidelines on the web, and 121 Japanese general physicians responded. Regarding the purpose of using the guidelines, their responses were as follows: confirmation of the therapeutic plan 86.8%, instruction for patients 38.0%, confirmation of diagnostic methods 35.5%. Their responses regarding the impact of the guidelines on clinical practice were as follows: major changes 0.8%, slight changes 57.0%, few changes 34.7%, no change 6.6%. As regards satisfaction, the results were: satisfied 3.3%, almost satisfied 55.4%, slightly dissatisfied 5.0%, dissatisfied 4.1%, none of the above 28.9%. The causes of dissatisfaction were difficulties in purchase 50.4%, insufficient citation of Japanese literature 31.4%, difficult applicability in clinical practice 23.1%. But, it should be stressed that the guidelines were written for stroke specialists and doctors or nurses working at institutes that have appropriate facilities for managing stroke patients.

In addition to the results of our study, these findings will be valuable in the continuous process of updating our guidelines, which is already ongoing under the guidance of the new Joint Committee on Renewing Guidelines for the Management of Stroke, in preparation for the Japanese Guidelines for the Management of Stroke 2009. Reflecting the above critical appraisals, we concluded that the updated guidelines should be improved in the following respects: (1) clarity of the recommendations, even in the absence of high-level evidence; (2) more consideration of patient preferences, applicability, implementation, and costs; (3) clear description of algorithms, potential benefits and harms, target users, and funding source/sponsor; (4) pilot study or external review before publication; (5) an English version, especially for other Asian countries; (6) version for patients and caregivers; (7) more comprehensive coverage, including transient ischemic attack, cerebral venous/sinus occlusion, asymptomatic cerebrovascular diseases, and specific causes of stroke, such as arterial dissection, Moyamoya disease, and so on; (8) improvement of availability of the book form; and (9) shortening of the interval between publication of revisions, and so on.

For the practice of evidence-based medicine in a true sense, physicians must consider not only published data, but also other factors that may be associated with a better prognosis. Therefore, physicians must consider each patient’s background including medical history, genetic inheritance, and financial and social status, as well as the ability of physicians, facilities available, and so on. Therefore, the guidelines should not be regarded as restricting each physician’s right of discretion, but rather should be regarded as a description of current standard ideas. Indeed, physicians are recognized as professionals because they have autonomous guidelines. Clinical practice guidelines do not become meaningful until they are formulated and used by physicians. Hence, the relationship between physicians and clinical practice guidelines is essentially interactive. The future of clinical practice guidelines depends on the professional awareness of physicians, who are always open to scrutiny by other physicians, patients, and the media. In this sense, we would greatly appreciate constructive criticism from all physicians and researchers who are involved in clinical practice or academic activities related to stroke.

Sources of Funding
This work was supported in part by a grant from the Japanese Ministry of Health, Labor, and Welfare, for Research on Health Technology Assessment.

Disclosures
None.

References
Postpublication External Review of the Japanese Guidelines for the Management of Stroke 2004

Yukito Shinohara, Masao Nagayama and Hideki Origasa

Stroke. 2009;40:1439-1443; originally published online February 19, 2009;
doi: 10.1161/STROKEAHA.108.535070

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2009 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/40/4/1439

An erratum has been published regarding this article. Please see the attached page for:
/content/40/9/e549.full.pdf

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Stroke is online at:
http://stroke.ahajournals.org/subscriptions/
Correction

In the article “Postpublication External Review of the Japanese Guidelines for the Management of Stroke 2004”, by Shinohara et al,1 several numerical values were given incorrectly, and the authors wish to make the following corrections, shown in italics. These corrections do not materially affect the results and have no effect at all on the conclusions of the article. The authors regret these errors.

1. On page 1439, in “Abstract”, in the 2nd line of the Results, “The average scores were 3.34, 3.53 and 3.51, respectively.” should be “The average scores were 2.98, 3.13 and 3.29, respectively”.

2. On page 1441, right column, line 4, “Among the unfavorable appraisals of the guidelines, ‘the guideline has been piloted among target users’ was common to all the subgroups (3.08, 3.00 and 3.17, respectively). Among the leading 5 items of the unfavorable appraisals, the items ‘the health benefits, side effects, and risks have been considered in formulating the recommendations’ (3.10) and ‘the potential organizational barriers in applying the recommendations have been discussed’ (3.11) were mentioned specifically by stroke specialists. The items ‘the recommendations are specific and unambiguous,’ ‘the patients’ views and preferences have been sought’, and ‘the target users of the guideline are clearly defined’ were mentioned specifically by nurse practitioners (3.18, 3.20, and 3.25, respectively).” should be “Among the unfavorable appraisals of the guidelines, ‘the guideline has been piloted among target users’ was common to all the subgroups (2.45, 2.38 and 2.27, respectively). Among the leading 5 items of the unfavorable appraisals, the items ‘the health benefits, side effects, and risks have been considered in formulating the recommendations’ (3.00) and ‘the potential organizational barriers in applying the recommendations have been discussed’ (2.91) were mentioned specifically by stroke specialists. The items ‘the recommendations are specific and unambiguous,’ ‘the patients’ views and preferences have been sought’, and ‘the target users of the guideline are clearly defined’ were mentioned specifically by nurse practitioners (3.18, 2.55, and 2.91, respectively).”

The corrected version can be viewed online at http://stroke.ahajournals.org.

1[Correction for Vol 40, Number 4, April 2009. Pages 1439–1443.]
(Stroke. 2009;40:e549.)
© 2009 American Heart Association, Inc.

Stroke is available at http://stroke.ahajournals.org

DOI: 10.1161/STROKEAHA.109.000022

e549