Standardization of the Clinical Diagnosis of the Dementia Syndrome and Its Subtypes in a Cross-National Study: The Ni-Hon-Sea Experience

Eric B. Larson,1 Susan M. McCurry,1 Amy B. Graves,2 James D. Bowen,1 Madeline M. Rice,1 Wayne C. McCormick,1 Nancy Zee,1 Akira Homma,3 Yukimichi Imai,4 Lon White,5 Kamal Masaki,5 Helen Petrovitch,5 Webb Ross,6 Michiko Yamada,7 Yasuyo Mimori,8 and Hideo Sasaki8

1University of Washington School of Medicine, Seattle.
2University of South Florida, Tampa.
3Tokyo Metropolitan Institute of Gerontology, Tokyo, Japan.
4St. Marianna University, Tokyo, Japan.
5Honolulu-Asia Aging Study, National Institute on Aging, Honolulu, Hawaii.
6Honolulu Veterans Affairs.
7Radiation Effects Research Foundation, Hiroshima City, Japan.
8Hiroshima University, Hiroshima City, Japan.

Background. Clinical investigators from Seattle, Honolulu, Tokyo, and Hiroshima participated in two standardization exercises in which data were collected on independent assessments. Exercises were conducted to evaluate the interobserver agreement on clinical diagnoses of dementia and dementia subtypes in a cross-national study of dementia prevalence and incidence rates in the United States and Japan.

Method. Fifteen clinicians from four participating sites assessed the diagnosis of 85 patients based on standardized summaries of clinical and diagnostic test data on each patient. Diagnostic guidelines and conventions were adopted on the basis of group consensus during standardization exercises.

Results. Using DSM-III-R criteria, generally good levels of agreement for all dementia diagnostic categories occurred in both years. For most measures of diagnostic agreement, improvements were observed between the 1995 and 1996 standardization sessions. Interrater agreement was highest for discrimination between dementia and non-dementia (1996 overall kappa, \( \kappa = .90 \)). The kappa values for dementia subtypes in 1996 ranged from .5 to .85, and for all sites combined the value was .67. For dementia subtypes, percent agreement was highest for vascular dementia and Alzheimer’s disease, but was less reliable for other types of dementia.

Conclusions. Clinicians from different cultures and medical traditions can reliably use the DSM-III-R criteria to classify dementia cases in cross-national research. The interrater agreement on dementia and its subtypes improved after clear-cut guidelines for interpretation of diagnostic criteria were developed and followed.

Migration of populations provides a unique opportunity to study genetic and environmental risk factors for chronic diseases. It also provides the opportunity to observe evolving care patterns in different cultures. When ethnic populations migrate into new cultures, they often experience relatively sudden changes in environment and lifestyle. Nonetheless, they tend to maintain the genetic homogeneity that distinguishes them from the majority population for at least several generations. Research projects focused on the relative influence of genetic and environmental factors or the interactive effects of multiple risk factors in ethnic groups can be enlightening, especially when the populations are studied in multiple locations across a historical migratory path.

A good example is the Ni-Hon-San study of Japanese migration from Japan (Ni) to Honolulu (Hon) and San Francisco (San). Over the years, a series of studies addressed the relationship between genetic and environmental factors, particularly in the development of coronary heart disease and cerebral vascular disease but also in the study of breast, colon, and stomach cancer. Historically, Japanese had the highest mortality rates from stroke among the developed countries and one of the lowest age-adjusted death rates from ischemic heart disease (1,2). However, early studies revealed that as Japanese persons migrated east to Hawaii and California, these disease patterns were altered such that stroke-related deaths were less common and coronary artery disease rates increased (3,4). Ulti-
mately, identification of environmental risk factors for stroke and heart disease in this genetically homogeneous group led to general behavior recommendations to reduce disease risk, including modifying dietary and smoking habits. Such cross-cultural studies also have provided valuable information on disease symptomatology, medical service utilization of minority and ethnic groups, and factors associated with successful aging (5).

Dementing illnesses present another interesting area for cross-cultural research in Japanese populations. Traditionally, vascular disease has been considered the most common cause of dementia in Japan, whereas Alzheimer's disease is the most common cause in the largely Caucasian populations of North America and Western Europe. Although overall dementia prevalence rates in Japan and the United States are similar, preliminary reports indicate that the composition of dementia subtypes is different. In Japanese studies, between 30% and 60% of dementia cases are attributed to vascular causes, and approximately half this level are attributed to Alzheimer's disease (6). By contrast, most European and North American populations are reported to have 50–70% of dementia cases due to Alzheimer's disease, whereas only 12–20% are attributed to vascular causes.

This set of observations led to the establishment of a study in Seattle (The Kame Project) of the prevalence and incidence of dementia and its subtypes in the Japanese American population of King County, Washington state. This study was planned and developed in collaboration with investigators in Honolulu (The Honolulu Aging Study), Hiroshima (Hiroshima Radiation Effects Research Foundation), and Tokyo, Japan. The goal of these studies, which are collectively called the Ni-Hon-Sea Study, is to identify whether rates and causes of dementia are the same or different across cultures. If the rates are different, the groups will be especially interested in the discovery of modifiable risk factors. The Ni-Hon-San study played a role in the identification of important risk factors for coronary heart disease and cerebral vascular disease and helped reduce the burden of these chronic diseases in both the United States and Japan. It is expected that identification of risk factors for dementia would likewise lead to the development of preventive measures to delay onset or to prevent its occurrence altogether.

For this research to be successful, however, standardization of procedures, particularly for case detection and diagnosis, is essential. One explanation of previously reported differences in disease rates has been different diagnostic criteria or differences in the application of the criteria themselves. This report describes the standardization efforts and results of the cross-cultural, cross-national study of elderly Japanese in Hiroshima, Tokyo (Ni), Honolulu (Hon), and Seattle (Sea). Specific details about the planning, organization, and conduct of individual participant studies are available elsewhere (7,8). We briefly describe the history of the project and then focus on the standardization process that evolved around case detection and diagnosis and the steps we used to resolve problems of disagreement. We also identify future areas of investigation that are needed to advance our understanding of cognitive function in aging societies around the world.

METHODS

Background

The University of Washington Alzheimer's Disease Research Center sponsored a visit of Japanese investigators to Seattle in September 1987. This one-day meeting included representatives from the Japanese American community in King County as well as scientific investigators; it was the beginning of a series of preliminary discussions concerning a collaborative study on dementia in the United States and Japan. In August 1988, the National Institute on Aging (NIA) released a request for applications for “cross-national investigations of the epidemiology of Alzheimer's disease and other dementias of later life” that would support research “in other countries, cultures, ethnic or population groups, with different exposures and habits . . . that may offer clues to the etiology of the disease that are not available here. The need to search more aggressively and widely for potent modifiable risk factors requires movement beyond national boundaries.”

Over the next several years, a series of international meetings including United States and Japanese investigators was held to develop a set of common research goals, standardized research methods, screening and diagnostic procedures, and plans for ultimate collaborative analyses. These efforts included development of the Cognitive Abilities Screening Instrument (CASI) (9), a 100-point measure intended for use with literate populations in cross-national studies of dementia. The CASI combines two of the most commonly used screening tests in the United States and Japan (the Mini-Mental State Examination and the Hasegawa Dementia Rating Scale). During this time period, it was administered and validated with nondemented elderly cohorts in Hiroshima, Honolulu, and Seattle (9,10).

In 1991, Seattle investigators received a 5-year NIA award to study the prevalence and incidence of dementia among elderly Japanese Americans in Seattle. The study was named the Kame Project after the Japanese word for turtle, a symbol representing longevity in Japanese culture. After numerous activities with the local Japanese American community and a county-wide census to identify Japanese Americans over age 55, a total of 1,985 Japanese Americans participated in a baseline examination and were enrolled (7).

Standardization Plan

Simultaneous with local efforts to launch the prevalence phase of the Kame Project, a standardization program was initiated to ensure that collaborating sites in Japan (Hiroshima and Tokyo) and the United States (Honolulu and Seattle) were collecting comparable and interpretable data. At a workshop in 1992, the core data collection procedures and measures were identified. Participants established a common set of research and standardization goals. Site visits to compare assessment and diagnostic procedures were conducted in Honolulu and at the sites in Japan in 1994. Standardization goals and steps that have been taken to achieve them are listed below for what has come to be known as the Ni-Hon-Sea study.

Standardization of forms, including screening tests, neuropsychological assessment, and risk factor data acquisition.—A core set of instruments was identified for use at all
sites, including the CASI for screening, the Consortium to Establish a Registry for Alzheimer’s Disease (CERAD) neuropsychological test battery (11), the Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) (12), and the Clinical Dementia Rating Scale (CDR) (13) to assess severity of dementia. In addition, sites followed CERAD clinical examination protocols, which specify the medical and family history, physical evaluation, and laboratory and neuroimaging tests (including head CT or MR) needed to derive a standardized dementia diagnosis (11,14,15). Dr. Homma and bilingual study interviewers from Seattle translated forms into Japanese. An independent translator provided back-translations for content comparisons between Japanese versions used at the various sites.

**Standardization of interviews.**—Cross-site visits were scheduled for clinical staff to allow observation of interviews, neuropsychological testing, and physical examinations. Training tapes in English and Japanese were created for the neuropsychological battery and physical examination.

**Analysis of CASI validity, reliability, sensitivity and specificity across sites.**—Because variable population characteristics, particularly age and level of education, could impact CASI scores, it was necessary to check performance of the instrument at each site. As expected, some calibration of cutoff scores was needed across sites to attain a similar level of sensitivity of the instrument to detect dementia cases, using clinical diagnosis by consensus committee as the "gold standard" against which the CASI was measured (9).

**Standardization of dementia diagnostic procedures.**—Researchers in the Ni-Hon-Sea studies participated in diagnostic consensus meetings during cross-site visits, providing an opportunity to review how clinical data were interpreted, differences in background data (e.g., imaging studies) that were available at each site, and which dementia subtypes were most problematic for achieving multisite agreement. The group decided that only moderate to severe cases of dementia would be counted, that is, cases in which a subject received a combined diagnosis of dementia using DSM-III-R criteria (16) and a dementia impairment severity rating (CDR) score of 1 or greater. This case detection approach recognizes the difficulty of reliable detection in early dementia cases and also minimizes the number of false positives diagnosed.

At the 1994 Japan and Honolulu site visits, participants developed a plan for a series of standardization workshops at which teams of investigators from all sites would meet to review cases and develop common diagnostic/interpretation criteria for problematic situations. Two such standardization workshops have been completed to date (1995 in Seattle, 1996 in Tokyo). Prior to both workshops, all sites were provided a package of subject records including clinical dictations, neuropsychological test data, and other pertinent data that had been used in the consensus diagnosis (such as CT scan films). Investigators at each site reviewed all records and arrived at a DSM-III-R diagnosis for each subject before convening with the larger group. For those diagnosed as demented, raters determined whether the cause of dementia was primary degenerative dementia of the Alzheimer’s type (AD), multi-infarct dementia (MID), or some other etiology. At each workshop, agreements and differences among sites were discussed, differences were resolved wherever possible, and diagnostic “rules” or conventions were established to minimize discrepancies in the future. Results from these two standardization meetings are summarized below.

**Statistical Methods**

For the 1995 standardization workshop in Seattle, a total of 55 subject records were reviewed by participants from each of the four primary sites (see Appendix A). In the 1995 exercise, all records were provided by the Seattle site. Records were selected in a stratified manner to provide a proportional representation of nondemented and demented (AD, MID, dementia, type other) subjects from the prevalence phase of the Kame Project. For the 1996 exercise, a total of 30 subjects were reviewed. Half of these were provided by Seattle and half by Honolulu. The Seattle site selected subjects at random in a stratified manner from the incidence phase of sampling, and the Honolulu site provided the first 15 subjects sampled for clinical workup. Subjects from the combined 85 charts ranged from 67 to 98 years (mean 82.2); 64% were women, and 27% lived in nursing homes or assisted living facilities.

Level of diagnostic agreement among the four rating sites was measured using the kappa statistic, which measures agreement over and above that occurring by chance, and percentage agreement statistics. Three subjects in 1995 and one in 1996 were not included in these computations because of missing diagnoses from one site. Kappa statistics were calculated to take into account multiple sites formulating multiple diagnoses (17) based upon four DSM-III-R groups (no dementia, AD, MID, unspecified dementia). Following the interpretation guidelines of Rosner (18), kappa values higher than .75 denote excellent interrater agreement, those from .40 to .75 denote good agreement, and those below .40 denote marginal agreement. Kappa statistics were compared using the normal approximation method (17).

**RESULTS**

Table 1 shows agreement among the four sites on the diagnosis of dementia syndrome (dementia/no dementia). For all combinations of rating sites, agreement improved between the two standardization years. All individual kappa indices reached statistical significance, and the overall kappa agreement among sites was significantly higher ($p < .0001$) in 1996 than in 1995 (Table 1).

Table 2 summarizes the intersite reliability of the clinical diagnosis of dementia subtype. When computing kappa statistics, less common dementia variants (e.g., dementia secondary to alcoholism or Parkinson’s disease), dementia cases with unknown etiology, and “mixed” dementia cases (most commonly those with a suspected combination of AD and vascular causes) were combined under the “dementia, type other” category. Mean kappa scores were lower than for dementia/no dementia comparisons, illustrating the increased difficulty in achieving agreement for diagnostic
Table 1: Interrater Agreements (Kappa Statistics) for DSM-III-R Diagnosis of Dementia, 1995 and 1996 Standardization Meetings*

<table>
<thead>
<tr>
<th>Site</th>
<th>1995†</th>
<th>1996†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995†</td>
<td>Site 2</td>
<td>.32</td>
</tr>
<tr>
<td>Site 3</td>
<td>.50</td>
<td>.46</td>
</tr>
<tr>
<td>Site 4</td>
<td>.28</td>
<td>.63</td>
</tr>
<tr>
<td>All sites combined</td>
<td>κ = .44, Z = 7.8, SE = .057</td>
<td></td>
</tr>
<tr>
<td>1996†</td>
<td>Site 2</td>
<td>.93</td>
</tr>
<tr>
<td>Site 3</td>
<td>.93</td>
<td>.86</td>
</tr>
<tr>
<td>Site 4</td>
<td>.93</td>
<td>.86</td>
</tr>
<tr>
<td>All sites combined‡</td>
<td>κ = .90, Z = 11.8, SE = .076</td>
<td></td>
</tr>
</tbody>
</table>

*Kappas based upon N = 52 in 1995 and N = 29 in 1996 due to missing diagnoses from one site.
†Two diagnostic categories: dementia versus nondementia.
‡Significantly different from 1995 at p < .0001.

Table 2: Interrater Reliability (Kappa Statistics) for Dementia Subtype, 1995 and 1996 Standardization Meetings, Using DSM-III-R Diagnostic Criteria*

<table>
<thead>
<tr>
<th>Site</th>
<th>1995 DSM-III-R†</th>
<th>1996 DSM-III-R†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995 DSM-III-R†</td>
<td>Site 2</td>
<td>.42</td>
</tr>
<tr>
<td>Site 3</td>
<td>.41</td>
<td>.38</td>
</tr>
<tr>
<td>Site 4</td>
<td>.49</td>
<td>.57</td>
</tr>
<tr>
<td>All sites combined</td>
<td>κ = .45, Z = 13.5, SE = .033</td>
<td></td>
</tr>
<tr>
<td>1996 DSM-III-R†</td>
<td>Site 2</td>
<td>.61</td>
</tr>
<tr>
<td>Site 3</td>
<td>.85</td>
<td>.52</td>
</tr>
<tr>
<td>Site 4</td>
<td>.79</td>
<td>.50</td>
</tr>
<tr>
<td>All sites combined‡</td>
<td>κ = .67, Z = 14.0, SE = .048</td>
<td></td>
</tr>
</tbody>
</table>

*Kappas based upon N = 52 in 1995 and N = 29 in 1996 due to missing diagnoses from one site.
†Four diagnostic categories: No dementia; Dementia of Alzheimer's type; Multi-infarct dementia; Dementia, type other.
‡Significantly different from 1995 at p < .0001.

Table 3: Interrater Agreements for All Sites on the DSM-III-R Clinical Diagnoses of Dementia

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All diagnoses combined</td>
<td>35</td>
<td>59</td>
<td>58</td>
<td>76</td>
</tr>
<tr>
<td>(N = 52 in 1995; N = 29 in 1996)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alzheimer's disease (AD)</td>
<td>25</td>
<td>33</td>
<td>56</td>
<td>78</td>
</tr>
<tr>
<td>(n = 12 in 1995; n = 6 in 1996)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multi-infarct dementia (MID)</td>
<td>46</td>
<td>60</td>
<td>67</td>
<td>87</td>
</tr>
<tr>
<td>(n = 13 in 1995; n = 5 in 1996)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dementia, type other</td>
<td>24</td>
<td>0</td>
<td>47</td>
<td>7</td>
</tr>
<tr>
<td>(n = 17 in 1995; n = 5 in 1996)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No dementia</td>
<td>50</td>
<td>92</td>
<td>67</td>
<td>97</td>
</tr>
<tr>
<td>(n = 10 in 1995; n = 13 in 1996)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Observed agreement based upon two diagnostic categories: yes versus no. For "gold standard" comparisons, percent agreements are based upon agreement of three sites against diagnosis given by the site that examined each patient and provided chart for review.
DISCUSSION

In this cross-national study of dementia diagnosis in elderly Japanese subjects, clinical investigators participated in a series of meetings designed to ensure that data were being collected in a comparable and interpretable way. The results of two standardization workshops (during which each site arrived at criteria-based diagnoses for 85 subjects based on review of standard medical record and neuropsychological test data) demonstrate that it is possible to achieve an acceptable level of agreement on clinical diagnosis in cross-national studies of dementia.

The sites achieved good kappa statistics and percent agreement levels (Tables 1–3). The level of agreement observed was consistent with data reported in other studies of cross-national, dementia interobserver agreement (19–21). Highest agreement levels were achieved with regard to the presence or absence of dementia. Agreement on subtypes varied, with the best agreement levels reported for vascular dementia and Alzheimer’s disease and the worst agreement levels for other diagnostic categories. Agreement levels improved in 1996 compared to the 1995 results, despite the fact that the 1996 exercise included incident cases of dementia, which are more difficult to classify because of the reduced historical information available for diagnosis as well as the fact that persons with a recent onset of dementia perform better on tests of cognitive and functional status than persons with more advanced disease. We believe this improvement reflects group learning from the previous consensus exercises, adherence to the diagnostic conventions developed during the 1995 meeting, and the shared knowledge the group gained of unique characteristics like education, lifestyle, and other features of the different cultures and sites.

Because individual differences between sites affected overall estimates of reliability, a goal of the standardization meetings was to clarify sources of disagreement. At closer examination, discrepancies between sites were often due to issues of data quality and interpretation. For example, critical information needed to make a DSM-III-R dementia diagnosis (e.g., documentation of remote memory loss or a decline in daily functioning) was sometimes omitted from the medical record. This was sometimes due to the fact that the 1996 exercise included incident cases of dementia, which are more difficult to classify because of the reduced historical information available for diagnosis as well as the fact that persons with a recent onset of dementia perform better on tests of cognitive and functional status than persons with more advanced disease. We believe this improvement reflects group learning from the previous consensus exercises, adherence to the diagnostic conventions developed during the 1995 meeting, and the shared knowledge the group gained of unique characteristics like education, lifestyle, and other features of the different cultures and sites.

Because individual differences between sites affected overall estimates of reliability, a goal of the standardization meetings was to clarify sources of disagreement. At closer examination, discrepancies between sites were often due to issues of data quality and interpretation. For example, critical information needed to make a DSM-III-R dementia diagnosis (e.g., documentation of remote memory loss or a decline in daily functioning) was sometimes omitted from the medical record. This was sometimes due to the fact that the 1996 exercise included incident cases of dementia, which are more difficult to classify because of the reduced historical information available for diagnosis as well as the fact that persons with a recent onset of dementia perform better on tests of cognitive and functional status than persons with more advanced disease. We believe this improvement reflects group learning from the previous consensus exercises, adherence to the diagnostic conventions developed during the 1995 meeting, and the shared knowledge the group gained of unique characteristics like education, lifestyle, and other features of the different cultures and sites.

Because individual differences between sites affected overall estimates of reliability, a goal of the standardization meetings was to clarify sources of disagreement. At closer examination, discrepancies between sites were often due to issues of data quality and interpretation. For example, critical information needed to make a DSM-III-R dementia diagnosis (e.g., documentation of remote memory loss or a decline in daily functioning) was sometimes omitted from the medical record. This was sometimes due to the fact that the 1996 exercise included incident cases of dementia, which are more difficult to classify because of the reduced historical information available for diagnosis as well as the fact that persons with a recent onset of dementia perform better on tests of cognitive and functional status than persons with more advanced disease. We believe this improvement reflects group learning from the previous consensus exercises, adherence to the diagnostic conventions developed during the 1995 meeting, and the shared knowledge the group gained of unique characteristics like education, lifestyle, and other features of the different cultures and sites.

Because individual differences between sites affected overall estimates of reliability, a goal of the standardization meetings was to clarify sources of disagreement. At closer examination, discrepancies between sites were often due to issues of data quality and interpretation. For example, critical information needed to make a DSM-III-R dementia diagnosis (e.g., documentation of remote memory loss or a decline in daily functioning) was sometimes omitted from the medical record. This was sometimes due to the fact that the 1996 exercise included incident cases of dementia, which are more difficult to classify because of the reduced historical information available for diagnosis as well as the fact that persons with a recent onset of dementia perform better on tests of cognitive and functional status than persons with more advanced disease. We believe this improvement reflects group learning from the previous consensus exercises, adherence to the diagnostic conventions developed during the 1995 meeting, and the shared knowledge the group gained of unique characteristics like education, lifestyle, and other features of the different cultures and sites.

Because individual differences between sites affected overall estimates of reliability, a goal of the standardization meetings was to clarify sources of disagreement. At closer examination, discrepancies between sites were often due to issues of data quality and interpretation. For example, critical information needed to make a DSM-III-R dementia diagnosis (e.g., documentation of remote memory loss or a decline in daily functioning) was sometimes omitted from the medical record. This was sometimes due to the fact that the 1996 exercise included incident cases of dementia, which are more difficult to classify because of the reduced historical information available for diagnosis as well as the fact that persons with a recent onset of dementia perform better on tests of cognitive and functional status than persons with more advanced disease. We believe this improvement reflects group learning from the previous consensus exercises, adherence to the diagnostic conventions developed during the 1995 meeting, and the shared knowledge the group gained of unique characteristics like education, lifestyle, and other features of the different cultures and sites.
ria and neuroimaging, including CT and MRI, will improve diagnostic rates, as will systematic comparison with autopsies. Improved understanding of the interrelationship between changes in brain circulation and the onset of illnesses such as Alzheimer’s disease will also enhance our clinical ability to classify persons as having vascular or Alzheimer’s type dementia.

Finally, because our standardization exercises to date have been based on review of materials from Seattle and Honolulu only, future work to demonstrate the diagnostic agreement across sites for Japan-based records will be needed to establish definitively that cross-national diagnostic reliability has been achieved. Nevertheless, our early findings are promising, and we expect that studies of aging and dementia in the Pacific Rim will grow. For example, since the Ni-Hon-Sea study was developed, investigators from Taiwan (see Appendix A) have adopted a similar methodology to study the prevalence and incidence of dementia in a largely rural population on the islet of Kinmen.

Conclusion

In summary, cross-cultural studies of the epidemiology of Alzheimer’s disease and related disorders require the development of standardized sampling, assessment, data handling, and diagnostic procedures. To be successful, such studies must also show that diagnostic outcomes are reliable across sites. The standardization exercise of the Ni-Hon-Sea study reported in this article demonstrates that investigators from different cultures can achieve acceptable levels of interrater agreement. Such agreement is important if we are to learn from the unique opportunity that migration of ethnic populations affords in understanding dementia and aging.

Acknowledgments

This study was supported in part by National Institute on Aging grants AG09769-03 and AG06781-06.

We thank Kame research coordinator Nancy Zee, as well as study interviewers and psychometrists Masaru Aoyama, Patti Taubota Boorkman, Lynne Salter, Megumi Sumitani, and Manami Wakuta. We also acknowledge the assistance of Drs. Brian Ito, Bud Kuukla, and David Nochlin, as well as Greta Hoshibata, Joyce Nakamura, Duane Beekly, Ross Barnhart, Steve Edland, Laura Gibbons, Mark Duprix, and Mary Jacka. Finally, we are grateful for the scientific contributions of Drs. Kazuo Hasegawa, Jim Mortimer, Dwayne Reed, and Evelyn Teng, and for the continuing support of our Community Advisory Board Members.

Address correspondence to Dr. Eric B. Larson, University of Washington Medical Center, Box 356330, Seattle, WA 98195-6330. E-mail: larson1@u.washington.edu

References


Received July 7, 1997
Accepted November 7, 1997
Appendix A

Participants at the 1995 and 1996 Ni-Hon-Sea International Standardization Meetings

Hiroshima Site: Drs. Michiko Yamada, Yasuyo Mimori, Hideo Sasaki, Junko Ikeda, Sinji Sudoh

Honolulu Site: Drs. Lon White, David Curb, Kamal Masaki, Helen Petrovitch, Webb Ross

Seattle Site: Drs. Eric Larson, Amy Graves, James Bowen, Wayne McCormick, Susan McCurry; Ms. Madeleine Rice, Nancy Zee

Tokyo Site: Drs. Akira Homma, Yukimichi Imai

Also attending

Taiwan Site: Drs. Hsiu-Chih Liu, Jon-Ling Fuh, Ching-Kuan Liu, Shuu-Jiun Wang

Appendix B

Sample Diagnostic Conventions Developed During 1995–96 Standardization Meetings of the Ni-Hon-Sea Study

- Only moderate dementia cases will be labeled as cases, i.e., subjects who meet DSM-III-R criteria for dementia plus have a Clinical Dementia Rating (CDR) score of 1 or greater. Subjects with early dementia (DSM+, CDR = 0.5) will not be considered cases regardless of DSM or NINCDS-ADRDA diagnosis.

- Sites may use any and all data that are available at consensus in determining a dementia diagnosis (including information from secondary sources, research studies, or specialized imaging tests).

- At standardization exercises, the diagnostic impressions of the examining on-site consensus team will be considered valid, since the greatest amount of clinical information is available to clinicians who actually examined each patient. For this reason, it is essential that sites include in dictated medical history and exam results information about disease onset and progression, cognitive impairment (particularly short- and long-term memory), functional status, and changes in personality or judgment.

- When conflicting information is provided by patient, informant, and physical exam, medical information will take priority. In a common example, if proxy states patient had a “stroke,” but there is no supporting documentation from medical records or physical exam, the reported stroke will not be included in diagnostic decisions.

Vascular Dementia

- A history of nonlocalizing neurologic symptoms (such as dizziness or sudden generalized weakness) will not qualify for a diagnosis of stroke or TIA, in the absence of other focal neurologic signs or laboratory evidence of cerebrovascular disease.

- Asymptomatic lacunar infarctions, periventricular lucency, or a history of hypertension will not be sufficient to make diagnosis of multi-infarct or vascular dementia.

- A single vascular event associated with a significant onset of decline in cognitive function is sufficient to diagnose multi-infarct dementia (stepwise decline), even in cases with otherwise insidious progressive course.

Miscellaneous Diagnostic Criteria

- Long-term memory is defined as ability to remember anything that is not “new” learning (including recent personal events such as yesterday’s dinner menu).

- Evidence of significant decline in social/occupational functioning is crucial to making a dementia diagnosis, and its presence or absence should be well documented in patients’ records. Very advanced age of dementia onset is not an exclusionary factor for AD.