Evaluation of Proxy Responses to the Stroke Impact Scale

Pamela W. Duncan, PhD, FAPTA; Sue Min Lai, PhD, MS, MBA; Denise Tyler, MA; Subashan Perera, PhD; Dean M. Reker, PhD, RN; Stephanie Studenski, MD, MPH

Background and Purpose—The purposes of this study were to compare proxy-patient responses on each domain of the Stroke Impact Scale (SIS) and the SIS-16, estimate the bias, and evaluate the validity of proxy scores.

Methods—Two hundred eighty-seven patient and proxy pairs from the Kansas City Stroke Registry participated in the study. All patients were assessed in their home or nursing facility between 90 and 120 days after stroke with the use of the modified Rankin Scale Motricity Index (strength), Barthel Index (activities of daily living), Lawton assessment (instrumental activities of daily living), Folstein Mini-Mental State Examination (cognition), and the SIS. Eligible proxies were individuals who were aged ≥18 years, had known the patient for at least 1 year, and saw the patient at least once each week. All proxy interviews were conducted within 7 days of (before or after) the patient’s interview.

Results—Three hundred seventy-seven patients from the Kansas City Stroke Registry were eligible for the study. Seventy-seven patients or proxies refused participation. Thirteen patients of the consenting patient-proxy pairs were too aphasic or cognitively impaired to complete the interviews and were dropped from the study. Proxies scored patients as more severely affected than patients scored themselves on the SIS-16 and in 7 of 8 domains of the full SIS (5 were statistically significant at α=0.05). The proxy bias toward overrating the severity of the patient’s condition tended to increase as the severity of the stroke increased. However, the magnitude of the biases between patient and proxy means, as measured by effect size, was small (range, −0.1 to 0.4). The strength of the agreement, as measured by intraclass correlation coefficients, between proxy and patient ranged from 0.50 to 0.83. Agreement was best for the observable physical domains. Both patient and proxy scores in all domains were significantly different across Rankin categories. Concurrent validity for both patient and proxy correlations with the Folstein Mini-Mental State Examination, Barthel Index, Lawton instrumental activities of daily living, and Motricity Index was good to excellent (range, 0.37 to 0.78).

Conclusions—Proxies provide valid information for assessment of stroke outcomes. There are significant differences between patient and proxy reporting on SIS domains and the SIS-16. However, the observed biases are small and not clinically meaningful. (Stroke. 2002;33:2593-2599.)

Key Words: disability evaluation ■ outcome ■ quality of life ■ stroke

The purpose of this study was to evaluate proxy-patient agreement on the Stroke Impact Scale (SIS). The SIS is a newly developed self-report outcome measure.1 The SIS Version 3.0 includes 59 items and assesses 8 domains (strength, hand function, activities of daily living [ADL]/instrumental activities of daily living [IADL], mobility, communication, emotion, memory and thinking, and social participation).1a Sixteen items from 4 of the 8 domains can be combined to produce a short composite physical domain score. The abbreviated physical domain is called the SIS-16.1b

One of the limitations of the use of self-report measures to assess outcomes after stroke is that many stroke survivors have cognition and communication problems. In a previous stroke outcome study, 25% of the subjects were excluded from health-related quality of life assessments because of cognitive and language disorders.2 In a large study that used mail-administered quality of life questionnaires, 50% of the stroke subjects were unable to complete the questionnaires by themselves.3 Study results can be seriously compromised and misleading if subjects who are suffering from severe deficits are excluded. The inclusion of proxy data will increase sample size, improve generalizability of the studies, and reduce sample bias.4

Good proxy-patient agreement is a necessary criterion for the usefulness of a stroke outcome measure.5 The development of the SIS requires that the reliability of proxy-patient responses to each domain be evaluated. The purposes of this study were to compare proxy-patient responses on each domain of the SIS and the SIS-16, estimate the biases, and evaluate the validity of patient and proxy scores. In addition, patient and proxy characteristics that affected agreement or bias were also evaluated.
Subjects and Methods

Recruitment of Subjects
All 300 patients taking part in this study were recruited through the Kansas City Stroke Registry. The Stroke Registry is funded by the National Institute of Aging Claude D. Pepper Older Americans Independence Center Grant and enrolls stroke patients from 17 Kansas City metropolitan area health facilities. Patients were enrolled if they were within 28 days of their stroke. Potential patients were identified in each of the 15 recruitment sites through a review of daily admission/discharge records and by referrals from unit-based nurses.

To be eligible for participation in the Stroke Registry, a patient had a documented diagnosis of stroke and ≥50 years. The World Health Organization definition of stroke was used for the purposes of this study. A stroke was defined as symptoms of rapid onset and of vascular origin reflecting a focal disturbance of cerebral function, excluding isolated impairment of higher function. These deficits must have persisted for ≥24 hours. In addition, patients could not have any of the following exclusion criteria: stroke onset ≥28 days, stroke due to subarachnoid hemorrhage, deficits from a previous stroke, not expected to live 1 year, New York Heart Association class IV heart failure, not community dwelling before stroke, not independent in basic ADL before stroke, living >50 miles from participating hospital, progressive or severe neurological disease, amputee, obtunded, comatose, or unable to follow a 3-step command.

If, on review of the medical record, a patient was found to be eligible for the Stroke Registry, consent to approach the patient about the Stroke Registry was obtained from the primary physician. The Stroke Registry was then explained to the patient, and informed consent was obtained.

Baseline Measures
At enrollment, a baseline assessment conducted by the research staff of nurses and physical therapists assessed patients’ demographics, stroke type, prior functional status, and stroke severity with the use of the National Institutes of Health Stroke Scale and Barthel Index.

Ninety-Day Follow-Up Assessments
All Stroke Registry patients received a follow-up telephone assessment at 90 days after stroke. Data were collected from the patient or a proxy. This telephone call consisted of a brief cognitive screen, the Hodkinson Cognitive Screen, to determine the patient’s ability to accurately answer questions about health status. If the patient was unable to communicate or score ≥5 on the Hodkinson Cognitive Screen, a proxy for the patient completed the telephone interview. During the telephone interview, the patient’s basic ADL and IADL were assessed with the Barthel Index and Lawton IADL. At the conclusion of this call, patients were asked to participate in an interview. Those who agreed were asked to identify a proxy. Eligible proxies were individuals who were aged ≥18 years, had known the patient for at least 1 year, and saw the patient at least once each week.

Patient Interview
All patients were assessed in their home or nursing facility by a master’s level research associate between 90 and 120 days after stroke. Patients’ 3-month stroke severity was characterized by the modified Rankin Scale. Strength was measured by the Motricity Index. The Folstein Mini-Mental State Examination (MMSE) was used to rate cognitive ability. After these measures were gathered, patients were interviewed with the use of the SIS.

Proxy Interview
All proxy interviews were conducted within 7 days of (before or after) the patient’s interview to minimize any actual change in the patient’s status between interviews. Both the patients and the proxies were blinded to each other’s responses. Before participation, each proxy signed an informed consent. Proxy demographics including age, sex, race, education, and relationship to the patient were gathered, as was information about the amount of time spent with the patient each week. Proxies were also asked to rate, on a scale of 0 to 100, how well they felt they knew the patient. Proxy mental status was then determined with the Folstein MMSE. Proxy health status was assessed by an interviewer-administered Medical Outcome Survey 36-item short-form health survey (SF-36). After this information was gathered, proxies were interviewed about patients’ stroke recovery with the SIS.

Data Analysis
Descriptive statistics were used to characterize the patients, proxies, and their relationships. Several analytical strategies were used to examine patient-proxy agreement. First, mean responses of the patients and proxies were compared for each of the individual SIS domains with the use of a paired-samples Student’s t test. To examine systematic bias, effect size was calculated by dividing the mean difference score by the standard deviation of the difference score. An effect size of 0.2 was considered a small bias, 0.5 a moderate bias, and 0.8 a large bias. Second, intraclass correlation coefficients (ICCs) between patient and proxy SIS assessments were computed with the use of the SAS VARCOMP procedure and minimum variance quadratic unbiased estimation. Guidelines used for the ICC as a measure of the strength of agreement were as follows: <0.40, poor agreement; 0.40 to 0.75, fair to good agreement; and 0.75 to 1.00, excellent agreement. Third, ANOVA was used to examine whether differences between proxy and patient assessments varied across the range of stroke severity (Rankin Scale 0 to 5). Finally, we used 17 regression models with interaction for each SIS domain to identify any patient and proxy characteristics that may have affected patient-proxy differences after controlling for the patient-proxy average SIS score.

To evaluate discriminant validity of patient and proxy scores, we examined differences in mean scores on each domain of the SIS across Rankin grades using ANOVA. We also assessed relationships between standard measures of stroke outcomes and select SIS domain scores with Pearson correlation coefficients. SAS version 8.0 was used for all statistical analyses.

Results
Three hundred seventy-seven patients from the Kansas City Stroke Registry were eligible to be included in this study. Fifty-seven patients refused to participate in additional assessments, and 20 proxies refused to respond to interviews. Three hundred stroke patients and their proxies took part in this study, representing 80% of all eligible patient and proxy pairs. There were no significant differences between those who refused and those who participated in terms of age (72.9 ± 8.9 and 72.8 ± 10.1 years, respectively; P = 0.9512), sex (53.3% and 53.3% female, respectively; P = 0.9892), or National Institutes of Health Stroke Scale score (6.6 ± 4.9 and 6.8 ± 4.7, respectively; P = 0.7384). Thirteen of the 300 patients who consented were too aphasic or cognitively impaired to complete the SIS interview (Table 1), and therefore this analysis is based on 287 patient and proxy pairs. Demographic characteristics of patients and proxies are shown in Table 2. The characteristics of the patients and proxies are presented in Tables 3 and 4.

Comparison of Patient and Proxy Responses
Agreement between patient and proxy SIS domain scores and the SIS-16 was examined in 287 patients for whom both patient and proxy information was available. Table 5 presents patient and proxy mean scores, patient-proxy differences, the magnitude of the differences (bias), and ICCs between patient and proxy reporting. Differences between patient and proxy
information were statistically different for 5 of the 8 SIS domains and for the SIS-16. In all 5 of these domains, the proxies rated the patients as more impaired than the patients rated themselves. However, the biases between patient and proxy means scores were low (0.1 to 0.4). The strength of the agreement between proxy and patient reporting ranged from moderate to excellent (ICCs ranged from 0.50 to 0.83). The best agreements were for domains that represented observable physical behaviors, and the worst agreements were on more subjective domains (eg, memory and thinking, communication, emotion, and strength).

Factors That Affect Patient-Proxy Differences
In patients with more severe stroke (Rankin Scale 4 to 5), the differences in patient-proxy reporting for ADL/IADL and SIS-16 were significantly larger than in individuals with less severe stroke. Stroke severity did not affect patient-proxy differences for the other SIS domains (Table 6). We used 17 regression models with interaction for each SIS domain to assess multiple patient and proxy characteristics (eg, age, sex, education, cognitive function, mental health, relationships, and time spent with the patients), which may affect patient-proxy differences. The results of these analyses did not systematically identify any patient or proxy factors that affected the observed differences.

Validity of Patient and Proxy Ratings
The relationships between patient and proxy SIS domains and some relevant standardized measures are reported in Table 7. The correlations between the Barthel Index and ADL/IADL domains and mobility domains of the SIS were very strong for both patient and proxy reporting. Agreement between Lawton IADL and SIS ADL/IADL domains was also excel-
The correlations between both proxy and patient reporting of memory functions and the Folstein MMSE were moderate (0.37 and 0.42, respectively). The mean scores for both patient and proxy reporting on all SIS domains and SIS-16 are depicted in Table 8. Both patient and proxy scores were significantly different across Rankin categories of stroke severity (0/1, 2, 3, and 4/5). The differences across Rankin categories were large for the physical domains but smaller for the memory, emotion, and communication domains.

**Discussion**

Although the necessity of including proxy responses to self-reported stroke outcome measures is clear, the reliance on proxy responses is valid only if the proxy responses are reliable and not compromised by systematic bias. The reliability of proxy responses has been assessed in studies of elderly subjects and in stroke survivors. Several findings are consistent across all studies of proxy respondents. First, proxy-patient ratings are more consistent when rating observable, concrete behaviors (eg, mobility and ADL). Agreement decreases when subjective judgment or estimation on behalf of the proxy (eg, questions pertaining to cognition or emotion or estimation of participation in activities outside the home) is required. Second, proxies tend to rate the subjects as more impaired or with a lesser quality of life than the subjects rate themselves. Third, lower levels of proxy agreement occur among severely affected subjects for whom proxy responses are the most needed.19,20

Agreement between stroke survivors and their proxies has been examined in 5 previous studies with the use of 6 different self-report measures. In a small study (n=110) of interviewer-administered assessments of chronic stroke survivors, Segal and Schall5 reported that proxy agreement was excellent (ICC=0.91) for the physical domain of the Functional Independence Measure but lower for the social-cognitive domain (ICC=0.61). Proxy agreement on IADL, as measured by the Frenchay Activities Index, was excellent (ICC=0.81) but was very poor for the SF-36 (average ICC=0.32; range, 0.67 for physical function to 0.15 for emotional role function).21 In a subsequent study of 25 of the subjects enrolled in the previous study, Segal and Schall5 compared proxy-patient agreement for telephone administration of the Functional Independence Measure. These subjects were assessed 18 months after stroke. Paralleling the results found for the in-person administration of the Functional Independence Measure at 6 months after stroke, the patient-
proxy agreement was 0.91 for the physical domain and was lower for the social-cognitive domain (0.52).

Sneeuw et al\textsuperscript{22} compared proxy-patient responses for a group of stroke survivors on the Sickness Impact Profile. ICCs were high (0.85) for the physical domains and moderate for the psychosocial domains (0.61). Systematic differences were observed between patient and proxy responses. Proxies rated patients as having more impairments than the patients rated themselves. The tendency of proxies to rate patients as more impaired was most pronounced in the more severely affected patients, but the bias was small. The proxy Sickness Impact Profile scores were significantly different across Rankin categories, which supports the validity of the proxy scores.\textsuperscript{22} Dorman et al\textsuperscript{23} compared proxy and patient responses for a group of stroke survivors on the EuroQol. They reported that proxy-patient agreements were best for self-care and physical function yet not much better than chance for psychosocial domains. Patients reported that their health states were better than the assessment of the proxies. Assessment of proxy-patient agreement on a telephone-administered community integration measure, the CHART, revealed that in a sample of 177 stroke patients the correlations were moderate to excellent for most domains but were lower for estimates of social integrations (ICC range, 0.44 to 0.72).\textsuperscript{24} The proxies rated themselves. The tendency of proxies to rate patients as having more impairments than the patients rated themselves. The proxy Sickness Impact Profile scores were significantly different across Rankin categories, which supports the validity of the proxy scores.\textsuperscript{22} Dorman et al\textsuperscript{23} compared proxy and patient responses for a group of stroke survivors on the EuroQol. They reported that proxy-patient agreements were best for self-care and physical function yet not much better than chance for psychosocial domains. Patients reported that their health states were better than the assessment of the proxies. Assessment of proxy-patient agreement on a telephone-administered community integration measure, the CHART, revealed that in a sample of 177 stroke patients the correlations were moderate to excellent for most domains but were lower for estimates of social integrations (ICC range, 0.44 to 0.72).\textsuperscript{24} The proxies were systematically more likely to rate the patients as more impaired, but the differences were small. The proxy-patient agreements for stroke survivors were similar to those for other disability groups.\textsuperscript{24}

The results of the current comparison of patient and proxy reporting on the SIS are consistent with previously reported studies in the elderly and in individuals with stroke.\textsuperscript{19–24} Proxy and patient reports on the SIS domains are significantly different, and proxies report the patients to be more impaired than the patients report themselves. However, the magnitude of the differences (bias) is small and not clinically meaningful. The average patient-proxy differences for SIS-16 range from \(-1.5\) to \(6.3\) (Table 5); however, the differences across Rankin severity levels for both patient and proxy reporting are large (Table 8). For example, the differences in scores for the SIS-16 physical domain across adjacent Rankin categories range from 11.9 to 25.5. The agreement between patient and proxy reporting is best for observable domains of ADL/IADL, mobility, hand function, and SIS-16. Agreements decrease when subjective judgment is required to assess strength, memory, emotion, communications, and social participation. The agreements for the subjective domains of the SIS are better than those reported for the subjective domains of the SF-36\textsuperscript{5} and the EuroQol\textsuperscript{23} and are equally as good as patient-proxy agreement on the social-cognitive domain of the Functional Independence Measure.\textsuperscript{5}

Similar to the findings of Sneeuw et al\textsuperscript{22} for the Sickness Impact Profile, a factor that affects patient-proxy agreement is stroke severity. In the most severely affected patients, the differences between patient-proxy reporting were greatest in the ADL/IADL and SIS-16 physical domains. However, the maximum average difference between patient-proxy reports (10.7 points), which occurred in the ADL/IADL domain, was less than half of the difference (22.8 points) between Rankin levels 3 and 4/5 for that domain.

Both patient and proxy ratings on the SIS are valid. Patient and proxy ratings are strongly correlated with clinician-administered standardized measures (Table 7). The weakest correlation was between the memory domain and the Folstein MMSE. This weak correlation may be due to 2 factors: (1) the low variance in Folstein MMSE scores and self-report memory deficits and (2) the fact that only 2 of the items of the Folstein MMSE capture memory. The between-level differences in Rankin Scale severity are significant for all domains of the SIS. As expected, given the emphasis of the Rankin Scale on physical function, the between-level differences were larger for the physical domains than for the memory, communication, and emotion domains.

### Table 7. Concurrent Validity: Pearson Correlations Between Selected SIS Domains and Other Assessments

<table>
<thead>
<tr>
<th>Assessment Domain</th>
<th>Patient Correlation</th>
<th>Proxy Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Folstein MMSE and SIS memory domain</td>
<td>0.42</td>
<td>0.37</td>
</tr>
<tr>
<td>Barthel Index and SIS ADL/IADL domain</td>
<td>0.72</td>
<td>0.78</td>
</tr>
<tr>
<td>Barthel Index and SIS mobility domain</td>
<td>0.69</td>
<td>0.70</td>
</tr>
<tr>
<td>Lawton IADL and SIS ADL/IADL domain</td>
<td>0.77</td>
<td>0.78</td>
</tr>
<tr>
<td>Motricity and SIS strength domain</td>
<td>0.67</td>
<td>0.69</td>
</tr>
</tbody>
</table>

# Conclusions

The evaluation of proxy-patient agreement on the SIS suggests that proxies may provide valid information for assess-
ment of stroke outcomes. This study provides further evidence that there are acceptable levels of proxy and patient agreement for observable physical behaviors and that the bias observed is not clinically meaningful. Clinicians and researchers should exercise more caution when proxy responses are used to assess the more subjective domains of emotion, memory, and communication. However, our estimates of agreement on the more subjective domains may be conservative because of the lower variability in scores in these domains in our sample.

The results of this study have implications for designing stroke outcome studies. It is appropriate to use proxy assessment; however, the researcher must be aware of the biases introduced with proxy responders. Proxies will report more limitations in function. Therefore, if an intervention improves outcomes and patients are more likely to be able to respond in the treatment group, the magnitude of the effect of the intervention may be exaggerated compared with the control group. This difference, however, may not affect the direction of the treatment effect.

Acknowledgments
This study was supported by the American Heart Association Pharmaceutical Roundtable Patient Care and Outcome Research Grant (9970089N) and the Claude D. Pepper Older Americans Independent Center Grant (AG-96-003). The following health facilities in the greater Kansas City area collaborated with us for patient recruitment: Baptist Medical Center, Bethany Hospital, Department of Veterans Affairs Medical Center at Kansas City, Independence Regional Health Center, John Knox Village Care Center, Liberty Hospital, Life Care Center of Grandview, Menorah Medical Center, Mid-America Rehabilitation Hospital, Overland Park Regional Medical Center, Rehabilitation Institute, Research Medical Center, Shawnee Mission Medical Center, Saint Luke’s Medical Center, Saint Joseph Health Center, Trinity Lutheran Hospital, and University of Kansas Medical Center.

References
Evaluation of Proxy Responses to the Stroke Impact Scale
Pamela W. Duncan, Sue Min Lai, Denise Tyler, Subashan Perera, Dean M. Reker and Stephanie Studenski

Stroke. 2002;33:2593-2599
doi: 10.1161/01.STR.0000034395.06874.3E

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/33/11/2593

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/