Validity of Proxies and Correction for Proxy Use When Evaluating Social Determinants of Health in Stroke Patients

Lesli E. Skolarus, MD; Brisa N. Sánchez, PhD; Lewis B. Morgenstern, MD; Nelda M. Garcia, BS; Melinda A. Smith, DrPH; Devin L. Brown, MD, MS; Lynda D. Lisabeth, PhD

Background and Purpose—The purpose of this study was to evaluate stroke patient–proxy agreement with respect to social determinants of health, including depression, optimism, and spirituality, and to explore approaches to minimize proxy-introduced bias.

Methods—Stroke patient–proxy pairs from the Brain Attack Surveillance in Corpus Christi Project were interviewed (n = 34). Evaluation of agreement between patient–proxy pairs included calculation of intraclass correlation coefficients, linear regression models (ProxyResponse = α0 + α1PatientResponse + β), where α0 = 0 and α1 = 1 denotes no bias) and κ statistics. Bias introduced by proxies was quantified with simulation studies. In the simulated data, we applied 4 approaches to estimate regression coefficients of stroke outcome social determinants of health associations when only proxy data were available for some patients: (1) substituting proxy responses in place of patient responses; (2) including an indicator variable for proxy use; (3) using regression calibration with external validation; and (4) internal validation.

Results—Agreement was fair for depression (intraclass correlation coefficient, 0.41) and optimism (intraclass correlation coefficient, 0.48) and moderate for spirituality (κ, 0.48 to 0.53). Responses of proxies were a biased measure of the patients’ responses for depression, with α1 = 4.88 (CI, 2.24 to 7.52) and α1 = 0.39 (CI, 0.09 to 0.69), and for optimism, with α1 = 3.82 (CI, −1.04 to 8.69) and α1 = 0.81 (CI, 0.41 to 1.22). Regression calibration with internal validation was the most accurate method to correct for proxy-induced bias.

Conclusion—Fair/moderate patient–proxy agreement was observed for social determinants of health. Stroke researchers who plan to study social determinants of health may consider performing validation studies so corrections for proxy use can be made. (Stroke. 2010;41:00-00.)

Key Words: stroke ■ proxy ■ social determinants of health
A convenience sample of consecutively interviewed ischemic stroke/transient ischemic attack patients were identified from September to November 2007 for this validation study. Eligibility was based on the ability of patients to correctly answer a brief set of questions to evaluate their cognitive and language capabilities. Eligible subjects were then asked to identify a proxy, defined as the person who knew the patient best, to participate in the validation study. Interviews took place at various intervals after the patients’ acute event based on patient availability and ease. Patients were queried in person (95%) or over the telephone (5%). Patients and proxies were asked identical questions, with proxies instructed to respond to the questions as they believed the patients would respond. For all patient–proxy pairs used in final data analysis, proxies were blinded to the patients’ responses. Proxy demographics and the type and duration of the proxy–patient relationship were collected. Written informed consent was obtained from all subjects, and the study was approved by the institutional review board at University of Michigan and the Nueces County hospital systems.

SDH Scales
The validation study focused on 3 measures of SDH: depression, optimism, and spirituality. Depression was measured with the Patient Health Questionnaire 9 (PHQ-9), which has been validated in stroke patients. For the purposes of this study, the PHQ-9 was scored as a continuous measure based on the total of the 9 individual questions (range, 0 to 27). Optimism was queried via a modified revised Life Orientation Test (LOT-R). Respondents indicated level of agreement with 6 statements on a Likert scale. Items from the optimism scale were presented both positively and negatively. Therefore, for the purposes of the analysis, some items were reverse scored. A composite optimism score was computed as the sum of the 6 individual responses, with a lower score corresponding to increased optimism (range, 6 to 24). Finally, spirituality was assessed using 2 ordinal spirituality questions was assessed by calculating percent agreement. Agreement between stroke patient and proxy responses ascertaining the importance of religious or spiritual beliefs was assessed using 2 questions to evaluate their cognitive and language capabilities. The intraclass correlation coefficient is a measure of overall agreement, ranging from 0 to 1, with higher values representing more agreement. Therefore, for the purposes of the analysis, some items were reverse scored. A composite optimism score was computed as the sum of the 6 individual responses, with a lower score corresponding to increased optimism (range, 6 to 24). Finally, spirituality was assessed using 2 questions to evaluate their cognitive and language capabilities.

Statistical Analysis
Descriptive statistics were calculated using medians and interquartile ranges (IQRs) for continuous variables. Categorical variables were analyzed using frequencies and percentages.

Evaluation of Agreement Between Stroke Patient and Proxy Responses for SDH
Agreement between stroke patient and proxy responses for the ordinal spirituality questions was assessed by calculating percentage agreement and weighted κ statistics. κ is a measure of overall agreement, ranging from 0 to 1, with higher values representing more agreement with values ≥0.5 representing fair to moderate agreement. Agreement between stroke patient and proxy responses for the continuous optimism and depression scores was assessed using intraclass correlation coefficients and linear regression models of the form: $\text{ProxyResponse} = \alpha_0 + \alpha_1 \text{PatientResponse} + \delta$. The intraclass correlation coefficient is a measure of overall agreement, ranging from 0 to 1, with higher values representing more agreement. The coefficients from the regression models describe the direction and extent of bias introduced by use of proxies (ie, when $\alpha_0 = 0$ and $\alpha_1 = 1$, the response of the proxy is unbiased).

Evaluation of Bias Introduced by Different Approaches to Incorporating Proxy Data
We conducted simulation studies to assess the extent of bias in future regression-based analysis of SDH–stroke outcome studies, which include proxy responses for SDH measures. The simulation study consisted of simulating data sets of 500 stroke patients according to assumed “true” models for the SDH–stroke outcome and patient–proxy associations, and subsequently estimating regression coefficients for the SDH–stroke outcome associations from the simulated data using the approaches described below.

### Table 1. Values for Factors in the Simulation Study

<table>
<thead>
<tr>
<th>Simulation Factor</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>1, 2</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.10, 0.25, 1, 2 (range of weak to strong effects)</td>
</tr>
<tr>
<td>SD of the error term, $\sigma$</td>
<td>2, 4, and 6</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>Point estimate in validation study, and lower and upper confidence limits</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>Point estimate in validation study, and lower and upper confidence limits</td>
</tr>
<tr>
<td>Standard deviation of the proxy’s response error, $\delta$</td>
<td>Observed SD in the validation study, as well as half and twice the observed SD</td>
</tr>
<tr>
<td>Percentage of patients with proxies</td>
<td>5%, 10%, 25%, and 50% of the sample</td>
</tr>
<tr>
<td>Total simulation scenarios per SDH scale</td>
<td>2592</td>
</tr>
</tbody>
</table>

For each data set, we first simulated patient responses to optimism and depression scales by resampling with replacement from the distributions observed in the validation study. We then simulated stroke outcomes for each patient assuming that the association between the outcome and optimism or depression followed a linear relationship (eg, $\text{Outcome} = \beta_0 + \beta_1 \text{Optimism} + e$). We then randomly selected a subgroup of patients as having only proxy responses. To generate proxy responses, we assumed a linear relationship: $\text{ProxyResponse} = \alpha_0 + \alpha_1 \text{PatientResponse} + \delta$ with $\alpha_0$ and $\alpha_1$ equal to those estimated with our validation study data, for both SDH scales. The simulation study followed a factorial design with 7 factors, with values shown in Table 1.

We estimated the regression coefficients $\beta_0$ and $\beta_1$ for the SDH–stroke outcome model, applying 4 approaches to the simulated data. First, we substituted the proxy responses in place of the patients’ responses and fit the regression model. The second approach was similar to the first, except we added an indicator variable for proxy use to the regression model (ie, adjustment for proxy use). Third, we used regression calibration, which consisted of substituting a “corrected” proxy response in place of the patient’s response (for patients with proxies). The corrected response was calculated as: $\text{CorrectedProxy} = \gamma_0 + \gamma_1 \text{ProxyResponse}$, where $\gamma_0$ and $\gamma_1$ were obtained from fitting a regression model $\text{PatientResponse} = \gamma_0 + \gamma_1 \text{ProxyResponse}$ to the validation data. This approach is termed “regression calibration with external validation” because $\gamma_0$ and $\gamma_1$ were obtained from our validation study and did not use any data from the simulated patient population. The fourth approach was similar to the third except that for each simulated data set, we simulated an additional 34 patients with proxies to serve as a validation sample and estimated $\gamma_0$ and $\gamma_1$ from that sample. This fourth approach is labeled “regression calibration with internal validation.” The 34 simulated internal validation patients were not used to estimate the outcome model so that the total number of patients remained at 500 across all simulated samples and so that the percentage of patients with proxies remained at 5%, 10%, 25%, and 50%.

For each combination of assumed SDH–stroke outcome and patient–proxy associations, 1000 data sets were simulated. For each data set, the percentage bias in the estimated $\beta_0$, $\beta_1$ was calculated as percentage bias = $100 \times (\hat{\beta}_0 - \beta_0)/\beta_0$. The percentage bias was averaged across the 1000 simulated data sets for the given scenario, and the corresponding 2.5th and 97.5th percentiles across the 1000 estimates of percent bias were calculated and were compared with 0.

As a way of summarizing the results of the 2592 simulation scenarios for each SDH measure, the relative importance of each simulation factor was assessed by conducting an ANOVA. For the ANOVA, the average percentage of bias estimated in each simulation scenario served as the outcome, and the simulation factors (values of various parameters) were explanatory variables.
Results

A total of 44 ischemic stroke/transient ischemic attack patients were eligible for this validation study. Patients were excluded if they refused to be interviewed separate from their proxy (n = 6), were unable to name a proxy (n = 3), or a proxy could not be located (n = 1). The validation study was completed with the remaining 34 patient-proxy pairs. Women comprised 59% of patients but 74% of proxies. Patients were 47% Mexican American whereas proxies were 44% Mexican American. Median age of patients was 63 years (IQR, 55 to 77) compared with 52 years (IQR, 40 to 63) for proxies. Proxies were most commonly the patients’ spouses (38%) or children (47%) and had long-term relationships with patients (median, 37 years; IQR, 25 to 55).

Median score on the PHQ-9 depression scale was 6 (IQR, 2 to 9), whereas median proxy score was 5 (IQR, 3 to 13). Fair agreement, intraclass correlation coefficient of 0.41, was found between patient and proxy responses. However, the proxies’ response was a biased measure of the patients’ response, with $\alpha_0 = 4.88$ (CI, 2.24 to 7.52) and $\alpha_j = 0.39$ (CI, 0.09 to 0.69) and residual SD of 5.22; ie, for patients with lower levels of depression, proxies overestimated the depression score by $\approx 5$ points, but this gap narrowed as the patients’ depression score increased. Median patient score on the optimism LOT-R scale was 11 (IQR, 9 to 14), whereas median proxy optimism score was 12 (IQR, 11 to 15). Fair agreement was also found between patient and proxy responses for optimism, with an intraclass correlation coefficient of 0.48. Again, the proxies’ response leaned toward being a biased measure of the patients’ response, with $\alpha_0 = 3.82$ (CI, $-1.04$ to 8.69) and $\alpha_j = 0.81$ (CI, 0.41 to 1.22) and residual SD of 2.57. Note that the CIs include 0 and 1, such that the statistical evidence of a biased proxy response is lower compared with the PHQ-9 scale. Furthermore, the residual SD is smaller, showing that although the proxies are somewhat biased in reporting optimism, their responses are generally more reliable than for depression.

Agreement for the spirituality questions was moderate, with $\kappa$ values of 0.55 and 0.46 and percentage agreement ranging from 74% to 79%. Notably, the majority of patients felt that spirituality was at least fairly important to what they do every day (88%).

Figure 1 shows the average percentage of bias and the percentiles in the estimated associations between a potential stroke outcome and the SDH measures obtained with various approaches to incorporating proxy data. Substituting the crude proxy responses resulted in biased regression coefficients, regardless of the use of a proxy indicator. Further, as the percentage of patients with proxy responses was increased, the bias increased. In contrast, the bias in the regression coefficients obtained with regression calibration with internal or external validation was nearly 0, although the precision of the estimates decreased as the percentage of patients with proxies increased. The estimated values of $\gamma_0$, $\gamma_j$ used to correct proxy responses in the regression calibration with internal validation varied across data sets; those used for regression calibration with external validation were $\hat{\gamma}_0 = 3.20$ and $\hat{\gamma}_j = 0.44$ for depression, and $\hat{\gamma}_0 = 6.17$ and $\hat{\gamma}_j = 0.41$ for optimism.

Figure 2 shows the impact of the variability in the proxy’s error, SD($\sigma$), on the percentage bias in the estimated SDH-stroke outcome association. When the variability of this error increases, the magnitude of the bias increases. Figure 3 shows the impact of the true association, $\beta_1$, when all other simulation factors remain fixed. As $\beta_1$ increases while the residual error SD, $\sigma$, remains fixed, the $R^2$ of the outcome model increases. The change in the $R^2$ ultimately impacts the precision of the estimated percentage bias. That is, when the $R^2$ is high (Figure 3, bottom), the bias in the estimated regression coefficient was negative for nearly every data set analyzed. However, when the $R^2$ is low (Figure 3, top), the...
degree of bias for a given data set is less certain and may be either positive or negative (although on average, it will be negative). The results were similar for optimism (data not shown).

Table 2 shows the relative importance of the simulation factors on the percentage of bias estimated in the simulation study. As can be observed, the percentage of patients with proxies and the SD of the proxy’s error are the most relevant factors impacting the bias in the estimated coefficients (largest mean squared error; \( P \) values <0.001 for all methods). The bias coefficients \( \alpha_0 \) and \( \alpha_1 \) explain relatively less variability in the percentage bias.

**Discussion**

This validation study found fair stroke patient–proxy agreement for the measurement of depression and optimism. Patient–proxy agreement was higher for spirituality but was nevertheless only modest. These results are consistent with previous studies that have shown greater stroke patient–proxy agreement in objective domains such as physical abilities but less agreement with more subjective domains such as energy, emotion, and mood. In this study, proxies commonly scored the patients as more depressed and less optimistic than the patients scored themselves. Other studies have also shown that proxies overestimate subjective end points such as poststroke quality of life and depression.

Given the potential importance of SDH to stroke outcomes, understanding methods to minimize bias in the use of proxy data is critical. We performed a simulation study to explore 4 different methods to correct for the proxy-induced bias. Our results suggest that directly substituting proxy data for the patient or adjusting for proxy use in models results in bias, and the degree of this bias increases with increasing frequency of proxy use and with decreasing reliability in the proxies’ responses. On the other hand, regression calibration with internal validation, where proxy data are corrected based on known (or estimated) information about the patient–proxy association, results in nearly unbiased estimates. Applying estimates of the patient–proxy association estimated with external validation studies reduced the bias compared with substituting proxy responses but did not always eliminate the bias (Figure 2). Extrapolating results from external validation studies warrants caution.

As an alternate to regression calibration methods, additional approaches that may improve the use of proxy information include developing more objective scales to access the patients’ SDH designed specifically for proxies. Studies have also proposed alternative statistical methods to correct for proxy-induced bias such as propensity score adjustment and psychometric profile analysis.

There are several limitations to this study. Because of the limited number of patient–proxy pairs, we were unable to compare agreement and bias in subgroups defined by gender, race/ethnicity, or relationship of the proxy to the patient. Further, the study population was limited to patients able to self-report their responses, and therefore the population...
studied is different from the population that actually requires proxies. Finally, also because of a small sample size, we were unable to assess whether there is lack of linearity in the association between proxy and patient responses. Linearity assumptions were used in the regression calibration methods, although nonlinear models can also be used to correct the proxy responses.

Proxies should continue to be included in stroke outcome studies to avoid selection bias. However, caution is recommended when collecting SDH data from proxies. In particular, appropriate methods should be used to incorporate proxy responses when there are larger percentages of patients with proxies and when the reliability of proxies is low. Researchers who plan to study SDH in stroke patients or other critically ill populations may consider performing validation work to identify and quantify the measurement error introduced by use of proxies in their population so corrections can be made. Additional research is needed to

Figure 3. Percentage bias in the estimated association between depression (PHQ-9) and stroke outcome when the true association is $\beta_1=0.1$ (top), $\beta_1=0.25$ (middle), or $\beta_1=2$ (bottom), but $\sigma$ remains constant at $\sigma=4$ (ie, outcome model $R^2=0.02$, 0.11, and 0.89, respectively), the SD of the proxy’s error is SD($\delta$)=5.2 (observed in the validation study), and varying the percentage of patients with proxy. Reg indicates regression; int, interior; ext, exterior; Inds, indicator variable for proxy use.

Table 2. ANOVA Results Using Percentage Bias in the Estimated Outcome–SDH Association as the Dependent Variable and Simulation Factors as Explanatory Variables

<table>
<thead>
<tr>
<th>Factor</th>
<th>df</th>
<th>Substitute Proxy</th>
<th>Substitute Proxy + Inds</th>
<th>Regression Calibration (Internal)</th>
<th>Regression Calibration (External)</th>
<th>Substitute Proxy</th>
<th>Substitute Proxy + Inds</th>
<th>Regression Calibration (Internal)</th>
<th>Regression Calibration (External)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>3</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>2</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.006</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>2</td>
<td>0.292*</td>
<td>0.000</td>
<td>0.000</td>
<td>0.015†</td>
<td>1.450*</td>
<td>0.000</td>
<td>0.001</td>
<td>0.312*</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>2</td>
<td>0.400*</td>
<td>0.593*</td>
<td>0.020*</td>
<td>0.700*</td>
<td>1.457*</td>
<td>0.107*</td>
<td>0.006*</td>
<td>0.521*</td>
</tr>
<tr>
<td>SD ($\delta$)</td>
<td>2</td>
<td>27.39*</td>
<td>31.22*</td>
<td>0.010*</td>
<td>5.583*</td>
<td>8.229*</td>
<td>18.91*</td>
<td>0.008*</td>
<td>3.30*</td>
</tr>
<tr>
<td>% Proxy</td>
<td>3</td>
<td>12.79*</td>
<td>10.80*</td>
<td>0.081*</td>
<td>0.926*</td>
<td>10.23*</td>
<td>7.49*</td>
<td>0.022*</td>
<td>2.06*</td>
</tr>
<tr>
<td>Residual‡</td>
<td></td>
<td>0.007</td>
<td>0.008</td>
<td>0.000</td>
<td>0.004</td>
<td>0.017</td>
<td>0.006</td>
<td>0.000</td>
<td>0.010</td>
</tr>
</tbody>
</table>

Degrees of freedom (df) and mean squared error associated with each simulation factor for each method of incorporating proxies and for each scale. Larger mean squared error implies that the factor explains a larger amount of variance in the percentage bias.

* $P<0.001$; † $P<0.05$; ‡ residual df=2576.
understand the bias introduced when other measures of SDH are under study.

Sources of Funding

This study was funded by the National Institutes of Health (National Institute on Neurological Disorders and Stroke; R01 NS38916). L.E.S. is funded by the American Academy of Neurology Clinical Research Training Fellowship. D.L.B. is funded by National Institute for Neurologic Disorders and Stroke (K23 NS051202). L.D.L. is funded by National Institute for Neurologic Disorders and Stroke (K23 NS050161).

Disclosures

None.

References

Validity of Proxies and Correction for Proxy Use When Evaluating Social Determinants of Health in Stroke Patients

Lesli E. Skolarus, Brisa N. Sánchez, Lewis B. Morgenstern, Nelda M. Garcia, Melinda A. Smith, Devin L. Brown and Lynda D. Lisabeth

Stroke. published online January 14, 2010; Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2010 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/early/2010/01/14/STROKEAHA.109.571703.citation

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/