Predicting Treatment Costs After Acute Ischemic Stroke on the Basis of Patient Characteristics at Presentation and Early Dysfunction

J. Jaime Caro, MDCM; Krista F. Huybrechts, MS; Heather E. Kelley, MA;
for the Stroke Economic Analysis Group*

Background—Given the pressure on healthcare budgets, assessing the cost of managing a disease has become a major research focus; yet collection of these data are labor intensive and difficult. Understanding the predictors of cost provides an efficient means of incorporating such information in decision-making concerning new therapies.

Methods—Data from two 12-week multinational trials that collected information on a variety of neurological, functional, and cost parameters for 1341 ischemic stroke patients were examined by means of multiple linear regression. Because the intent is for the model to be predictive, only patient characteristics that can be known at the time of patient presentation or shortly thereafter were evaluated for inclusion in the model.

Results—The Barthel Index was the strongest predictor of cost in all models evaluated. Other major predictors, either directly or through their impact on survival, were stroke subtype, neurological impairment, congestive heart failure, and country. A good model fit was obtained, judging by the model statistics (model $F=84$, 3 *df*, $P<0.0001$) and the accuracy of the predictions (<3% difference between mean actual and predicted cost).

Conclusions—Through the use of key patient characteristics, this regression model allows for prediction of the cost of stroke care, which may be helpful in the context of therapeutic decisions and budgetary planning purposes. It also provides insight into how specific treatments, through their impact on clinical characteristics, can modify the cost of stroke treatment. **(*Stroke*. 2001;32:100-106.)**

Key Words: outcome ■ statistical models ■ stroke ■ stroke management

Because most industrialized countries have reached the point where healthcare costs are threatening to exceed what they can reasonably afford, regulators are adopting a rigorous attitude toward the acceptance and dissemination of new, often more expensive, therapies. This evolution has led to a heightened interest on the part of all stakeholders in understanding the cost of care, its determinants, and how they may be modified by specific treatments; note the recent growth in dedicated cost studies across therapeutic areas.[1,2] To address these concerns, patient healthcare resource utilization was documented extensively in the context of 2 phase III clinical trials undertaken to establish the efficacy and safety of a potential new neuroprotective agent.[3,4] Although the clinical development program was discontinued, the studies provide an invaluable source of information on numerous aspects of stroke management across settings and its associated costs.

Descriptive analyses of the intensity of health service use by various patient and disease characteristics have been reported previously.[5] The mean cost of acute stroke management during the 12-week period studied amounted to 13 668 US dollars (USD); which is equivalent to the cost of ≈46 hospital days. More than 70% of this cost was accounted for by the initial hospitalization, which averaged 24 days. The total cost and its components varied according to patient age, the presence of comorbidities, and several indicators of disease severity, including functional and neurological impairment and stroke subtype. The purpose of the current analyses was to seek combinations of these factors that predict the costs of managing stroke over the first 3 months.

Subjects and Methods

Study Description

The trial procedures and the clinical and economic outcomes considered have been described in detail elsewhere.[3–5] In summary, a total of 1446 stroke patients, ≥18 years of age who presented with substantial neurological deficit (European Stroke Scale [ESS] <70 or National Institutes of Health Stroke Scale [NIHSS] >7) within 6 hours after stroke onset, were enrolled in the trials. One trial was conducted in North

Received June 21, 2000; final revision received September 12, 2000; accepted September 12, 2000.

From Caro Research, Concord, Mass (J.J.C., K.F.H., H.E.K.); and the Division of General Internal Medicine, Royal Victoria Hospital, McGill University, Montreal, Quebec, Canada (J.J.C.).

*Members of the Stroke Economic Analysis Group are listed in the Appendix.

Correspondence to J. Jaime Caro, Caro Research, 336 Baker Ave, Concord, MA 01742. E-mail jcaro@caroresearch.com

Stroke **is available at http://www.strokeaha.org**

America (United States and Canada), the other in Australia and Europe (Austria, Belgium, Denmark, Finland, France, Germany, the Netherlands, Norway, Sweden, and the United Kingdom). The protocols for both studies were virtually identical and were approved by the relevant institutional review boards. All subjects provided informed consent. Patients received active treatment or placebo for 5 days or until complete neurological recovery, whichever occurred first, and their progress was monitored for a period of 3 months after treatment initiation. Neurological recovery was evaluated on the basis of the NIHSS in the North American study and the ESS in the international study. Functional recovery was measured by means of the Barthel Index and Rankin Scale in both studies.

Detailed information on inpatient and community healthcare use was collected during the trial period. Inpatient management covered the time spent in either acute or long-term care facilities. Outpatient care encompassed various types of rehabilitation therapy and other healthcare services patients received while they stayed at home or in a retirement home. Modifications to patients' homes and medical equipment purchases made to mitigate the disability caused by the stroke were also documented. A comprehensive description of the approach used to collect the resource use information has been provided elsewhere.[5]

Unit costs (1996) from the United Kingdom[6,7] were used to aggregate the different types of resource use across countries and permit estimation of the overall burden of stroke care. Aggregation of resource use across countries is a complex and controversial issue, and its implementation is a topic of extensive debate among health economists.[8] The rationale for and the strengths and weaknesses of the particular approach adopted in this study have been discussed previously.[5] In essence, it was believed that use of the local unit costs from each participating country would artificially create cost differences attributable to differences in reimbursement systems and cost structure, as opposed to true differences in the underlying management patterns, and would, as such, unnecessarily obfuscate the findings. Because the study is conducted from the perspective of the healthcare system, costs that are not reimbursed by a government insurance plan are not included in the analyses. To facilitate interpretation of the findings by an international audience, results are reported in US dollars, with an exchange rate–based conversion (1 USD=0.61 £).

Analyses

Multiple linear regression analyses were carried out to determine the predictors of the total 12-week treatment cost. Because the intent is for the equations to be predictive, we evaluated only those patient characteristics that can be known on the basis of the examination at the time of first presentation to the emergency room or physician's office or in the first few days immediately following. These characteristics were age (continuous); sex; country of treatment (United States or non–United States); living situation before the stroke (alone or living with a partner); Barthel Index before the stroke (on a 0 to 100 scale, categorized in 4 levels: 0 to 45, 50 to 70, 75 to 95, 100); comorbidities (any present versus absent); physician's clinical global impression at time of presentation (mild, moderate, or severe stroke); stroke subtype (small-vessel occlusive, large-vessel occlusive, cardioembolic stroke, or stroke of undetermined cause, based on the TOAST criteria[9]); Barthel Index in the first 5 days after stroke (continuously on a 0 to 100 scale); neurological deficit assessed on the basis of the ESS or the NIHSS (5 severity levels were defined: very mild [NIHSS: 0 to 9, ESS: 60 to 100], mild [NIHSS: 10 to 12, ESS: 46 to 59], moderate [NIHSS: 13 to 15, ESS: 38 to 45], severe [NIHSS: 16 to 19, ESS: 28 to 37], and very severe [NIHSS: 20 to 34, ESS: 0 to 27]); and finally, admission to a stroke unit during the initial hospitalization. The latter was considered of interest because some studies have found that use of a stroke unit may be associated with expedient discharge from the hospital to a nursing home or home as the result of a timely and multidisciplinary integrated treatment approach.[10–12] Stepwise variable selection was used to identify the most significant predictors. All variables of interest, prompted by the univariate analyses—which have been described previously[5]—together with clinical credibility, were entered initially. The statistical criterion for considering retention of a variable in the equation was $P<0.05$. Before selection of the final equations, the subsets of variables retained were evaluated on their clinical and predictive relevance. The strength of the predictive ability of the variables retained was assessed by forming bootstrap samples and refitting the model in each sample.

Death has a paradoxical impact on the total treatment cost because, although the cost per surviving day among the patients who die during the follow-up period tends to be higher, the abbreviated time over which costs accumulate tends to make the total cost lower.[5] Any model predicting the total management cost will thus have to account explicitly for this effect. Because the patient's eventual vital status is not known early on—a requirement for all covariates entered in the model—a 2-part model was developed. First, a logistic regression analysis was carried out to provide the probability that the patient will die during the course of the 12 weeks. This probability is then entered as a potential predictor in the multiple regression analysis predicting total treatment cost.

A split-sample approach was used to evaluate the reliability of the equations. Two thirds of the patients were randomly assigned to the training group and one third to the validation group. The best-fitting equations were developed by using the data for the training group and subsequently were validated by applying them to the patients in the second sample. Specifically, the model reliability was assessed by comparing the mean predicted cost with the actual mean cost in the validation group.[13]

As is frequently the case with medical cost data, the distribution of the 12-week treatment cost is positively skewed.[14] The values were therefore logarithmically transformed to achieve a more normal distribution and permit use of standard parametric statistical tests. The equations were derived with the use of these log costs.

Because the aim is to predict costs rather than log costs, the predicted log costs must be transformed back. Simply calculating the predicted log costs and then exponentiating them

TABLE 1. Characteristics of 1341 Patients Included in Analyses

Demographics	
Age, y (mean, SD)	70.5 (12.8)
Sex, male	55.0%
Country	
United States	44.2%
Non–United States	55.8%
Living situation before stroke*	
Alone	36.7%
Living with a partner	63.3%
Comorbidity, any	85.4%
Hypertension	55.9%
Coronary artery disease	36.6%
History of atrial fibrillation	28.0%
Diabetes	20.7%
Previous myocardial infarction	19.2%
Congestive heart failure	17.0%
Previous TIA	10.2%
Previous stroke	7.8%
Renal problems	4.1%
Hepatic problems	1.7%
Barthel Index before stroke†	
0–45	0.9%
50–70	2.2%
75–95	14.0%
100	82.9%

*Living situation unknown for 32 patients.
†Information missing for 4 patients.

will introduce bias because of the nonlinearity of the transformation.[14] Instead, the appropriate retransformed estimates are obtained by multiplying the exponential of the individual patients' predicted log cost with a smearing estimator, defined as the average of the exponentiated residuals.[15]

Standard regression diagnostics were used to check the assumptions for the regression analyses.[13] Analyses were carried out with SAS version 6.12 for Windows.

Results

A total of 105 patients were excluded from the analyses: economic information was missing for 34 patients, and another 71 did not have an ischemic stroke. The demographics and clinical status on presentation for the 1341 patients included in the analyses are summarized in Tables 1 and 2.

Model Step 1: Vital Status

With the use of logistic regression analyses, death during the first 12 weeks after stroke was found to be predicted (-2 log likelihood $\chi^2=235$, 5 *df*, $P<0.0001$) by physical disability in the 5 days immediately after presentation (according to the Barthel Index), a clinical history of congestive heart failure, very severe neurological impairment (based on the ESS or NIHSS), a large-vessel occlusion, and whether the country was the United States. The resulting equation is presented in Table 3. From it, the probability of dying in the first 12 weeks after a stroke can be estimated.

Model Step 2: Total Management Cost

The best-fitting equation of those examined (model $F=84$, 3 *df*, $P<0.0001$) contained 3 factors. Early functional disability as reflected by the Barthel Index was the strongest predictor of the 12-week management cost, followed by the probability of dying. One other factor contributed significantly to the prediction of the total management cost: cardioembolic stroke. The Barthel Index remained the strongest determinant regardless of the order in which the variables were entered in the model, and in all but 1 of the 100 bootstrap samples, it entered first.

Although physical functioning was clearly anticipated to be a predictor, the strength of its association with cost, relative to other factors, is quite striking. When ignoring the impact of the Barthel Index on death, a 5-point change in the Barthel Index causes an absolute change in the log cost of 0.095, which corresponds to a change of ≈10% in the retransformed cost. Stroke subtype is also a strong predictor of cost, with a cardioembolic stroke having a 15% higher cost than the small- and large-vessel occlusions. This finding indicates that stroke subtype carries prognostic implications beyond severity and dysfunction. Severity of neurological impairment as measured by the ESS or NIHSS was important in establishing the probability of death but did not add predictive power beyond that already contributed by the

TABLE 2. Clinical Status on Presentation for 1341 Patients Included in Analyses

Clinical Status on Presentation	
Stroke subtype*	
Small-vessel occlusion	8.4%
Large-vessel occlusion	37.7%
Cardioembolic	28.3%
Clinical global impression†	
Mild	4.4%
Moderate	40.2%
Severe	55.4%
Neurological impairment (ESS/NIH)†	
Very severe	21.3%
Severe	22.8%
Moderate	17.9%
Mild	20.4%
Very mild	17.5%
5-d Barthel Index‡	
0–45	75.1%
50–70	9.7%
75–95	7.5%
100	7.8%
Admitted to stroke unit	44.5%

*Type undetermined for 343 patients (25.6%).
†Information missing for 1 patient.
‡Seventy-two patients died before evaluation of the Barthel Index and were assigned a score of 0.

TABLE 3. Logistic Regression Analysis Predicting Death During First 12 Weeks After Stroke Onset in Training Group (n=894)

Variable	β (SE)	P
Intercept	−0.631 (0.218)	0.0037
Barthel Index	−0.080 (0.010)	<0.0001
Congestive heart failure	0.724 (0.241)	0.0027
Severe neurological impairment	0.440 (0.209)	0.0354
Large-vessel occlusion	−0.441 (0.211)	0.0369
Country	0.484 (0.201)	0.0161

Barthel Index: continuous variable based on scale from 0 to 100; congestive heart failure: yes=1, no=0; severe neurological impairment: NIHSS 20–34, ESS 0–27; large-vessel occlusion: yes=1, no=0; country: United States=1, non–United States=0. −2 log likelihood χ^2=235.171 (5 *df*), *P*<0.0001.

Barthel Index. It only entered in 12% of the bootstrap samples and never ahead of the Barthel Index. Thus, it appears that the relation between neurological impairment and cost is mediated by its association with other predictors. Given that two different instruments were used to measure neurological impairment in the trials, the ESS and NIHSS, neurological deficit was entered in the equations as a categorical rather than continuous variable. This may have contributed to the more limited explanatory power of this factor when compared with the Barthel Index. Age and the presence of comorbidities per se were not significant additional predictors of cost. Individual comorbidities were looked at separately and were occasionally but inconsistently predictive, depending on the model (data not reported).

To increase its usefulness, an expanded equation that includes demographic characteristics—in particular, age, sex, and country—was also derived. Both this full and the reduced model (including only statistically significant predictors) are summarized in Table 4.

To evaluate the model reliability, we applied this 2-part model to the validation sample consisting of the remaining 447 patients. The actual mean cost in this group amounts to 13 107 USD. Because the reduced model overestimates this true cost by <10%, the model was judged to be reliable.

Retransformation

To meet the objective of predicting cost rather than log cost, smearing estimators were incorporated in the retransformation: 1.283 for the reduced model and 1.281 for the full model. Exploration of the residual errors, however, indicated the presence of continued heteroscedasticity by Barthel Index categories, in which case use of an overall smearing estimator provides estimates of the expected cost that are still somewhat incorrect.[16] Separate smearing estimators were therefore calculated for 3 subgroups of patients, defined according to their Barthel Index. The smearing estimators for these subgroups, as well as a comparison of the mean actual and the mean retransformed predicted costs, are presented in Table 5. The difference between the actual and predicted costs is ≈3% for the overall population and varies between ≈0% and 5% for the subgroups, when using the specific smearing factors.

Example

As an example of how to compute the predicted 12-week cost, consider an 82-year-old woman with a cardioembolic stroke who lives in the United States and has an early Barthel Index of 65 and an NIHSS score of 17. By using the coefficients from the logistic regression model in Table 3, begin by computing the log odds that this patient will not survive the first 12 weeks after stroke: $-0.631+(-0.080\times65)+(0.724\times0)+(0.440\times0)+(-0.441\times0)+(0.484\times1)=-5.347$. The probability of dying is thus derived by using $1/(1+e^{-(-5.347)})=0.005$. Next, compute the predicted log-cost based on the full model specified in Table 4: $10.350+(-0.019\times65)+(-2.603\times0.005)+(0.143\times1)+(-0.001\times82)+(0.037\times0)+(-0.050\times1)=9.114$. The retransformed predicted 12-week cost for this patient equals $e^{9.114}$x1.212 (smearing estimator)=11 003 in 1996 USD. The mean cost for a specific patient population would be obtained by repeating this calculation process for each individual patient in the population and then averaging the results (it is inappropriate to use the mean population values in the equations to directly compute a "mean" cost).

Because the Barthel Index affects both the probability of dying and the management cost for a given survival status—as

TABLE 4. Multiple Linear Regression Analysis Predicting Log-Transformed 12-Week Management Cost in Training Group (n=894)

	Reduced Model		Full Model	
Variable	β (SE)	P	β (SE)	P
Intercept	10.307 (0.075)	<0.0001	10.350 (0.177)	<0.0001
Barthel Index	−0.019 (0.001)	<0.0001	−0.019 (0.001)	<0.0001
Probability of dying	−2.674 (0.211)	<0.0001	−2.603 (0.220)	<0.0001
Cardioembolism	0.137 (0.060)	0.0232	0.143 (0.061)	0.0192
Age			−0.001 (0.002)	0.6828
Sex			0.037 (0.053)	0.4844
Country			−0.050 (0.054)	0.3603

The reduced model only includes significant predictors. The full model controls for demographic characteristics. Barthel Index: continuous variable based on scale from 0 to 100; probability of dying: estimated based on logistic regression model (see Table 3); cardioembolism: yes=1, no=0; age: continuous; sex: male=1, female=0; country: United States=1, non–United States=0. The model *F* for the full model is 42 (6 *df*), *P*<0.0001; for the reduced model, 84 (3 *df*), *P*<0.0001. The adjusted R^2 equals 0.22 for both models.

TABLE 5. Comparison of Mean Actual and Retransformed Predicted Costs in the Training Group for Reduced and Full Models, With Overall Smearing Estimator and Separate Smearing Estimators by Subgroup

	Overall	Barthel Index 0	Barthel Index 5–70	Barthel Index ≥75
No. of patients	894	217	534	143
Actual cost (USD), mean (SD)	13 949	14 060	15 875	6587
	(9 852)	(13 635)	(7 824)	(5415)
Reduced model				
Smearing estimators	1.283	1.482	1.209	1.260
Retransformed predicted cost (USD), mean (SD)*	14 354 (1 269)	14 669 (1 717)	16 087 (1 358)	6596 (777)
% Difference actual vs predicted	2.90%	4.33%	1.33%	0.13%
Full model				
Smearing estimators	1.281	1.472	1.212	1.253
Retransformed predicted cost (USD), mean (SD)*	14 343 (3 517)	14 684 (3 844)	16 101 (3 911)	6581 (1685)
% Difference actual vs predicted	2.82%	4.44%	1.42%	−0.09%

*SD calculated by bootstrapping approach.[24]

can be deduced from the 2-step model—the effect of changes in the Barthel Index on the total management cost is nonlinear, as illustrated in the Figure for a few sample patients.

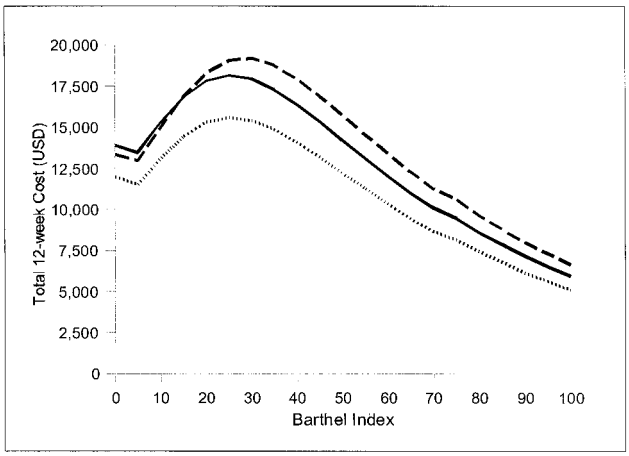
Discussion

Through the analysis of 2 large clinical trial data sets that contained extensive information on both clinical and economic outcomes, a 2-part multiple regression model has been developed that allows for prediction of the overall 12-week stroke management cost, making use of key demographic and disease characteristics known at the time of or shortly after patient presentation. The Barthel Index was the strongest predictor of cost in all models evaluated. In the final model presented here, a 5-point difference in the Barthel score (on a scale from 0 to 100) predicts a decrease in the total cost of ≈10%. A diagnosis of cardioembolic stroke was another significant predictor.

Several other factors acted more through their effect on the probability of dying. For example, although severe neurological impairment can be a direct predictor of cost, its effect is better captured as a significant determinant of stroke death. It should be noted, however, that identification of the causal factors and quantification of their impact on mortality was not a specific purpose of this study. Estimation of the 12-week mortality is only an intermediate step required to properly account for the length of time during which costs are incurred.

Many analyses examining the determinants of stroke cost and/or clinical outcomes focus on relatively small and homogeneous patient populations during one segment of the continuum of care, be it acute care, inpatient rehabilitation, or extended rehabilitation in a community setting. These studies are particularly useful to individual institutions or practitioners who are charged with the responsibility of reducing the costs while maintaining the quality of care in that specific segment. Our analysis moves away from this care segmentation and takes a more global approach by defining a set of predictors of stroke cost across the entire continuum of care, within the restrictions of the 12-week study period.

Comparison of specific findings across studies is difficult because of the vast differences in objectives, patient populations studied, parameters evaluated, measurement instruments used, and analytical approach. Broadly speaking, however, our findings are in line with previous reports on the determinants of stroke costs, although the role of physical functioning tends to be more pronounced in our analyses. Harvey and colleagues,[17] for instance, concluded that severe neurological impairment (NIHSS-based) and physical disability (based on the motor score of the Functional Independence Measure) are the main predictors of longer length of



Barthel Index affects both probability of dying and management cost for given survival status, as can be deduced from 2-step model. Effect of changes in Barthel Index on total management cost is nonlinear. Solid line indicates 82-year-old woman with cardioembolic stroke, living in the US, with an NIHSS score of 17; dotted line, 90-year-old woman with small-vessel stroke, living in US, NIHSS=10; and broken line, 55-year-old man with cardioembolic stroke, living outside US, ESS=50.

stay in a rehabilitation setting. Research by Galski and colleagues[18] has shown that mainly the higher-order cognitive impairments extended the length of stay and increased the referrals for outpatient therapies and home services after discharge. The Copenhagen Stroke Study reported that longer inpatient length of stay, encompassing both the acute and rehabilitative care, was driven by stroke severity (evaluated by the Scandinavian Stroke Scale), marital status, and death.[12] In that respect, it is interesting to note that living situation did not emerge as a predictor of total cost in our study.

To our knowledge, this study is the first in which the role of the Barthel Index as a predictor of cost, and thus resource use intensity, is so manifest and, in fact, stronger than the role of neurological impairment. This is an encouraging finding in light of the fact that the Barthel Index is a well-researched instrument that is easy to administer and has been found to be reliable, valid, and sensitive.[19]

The predictive equations should be particularly useful in the context of economic modeling of treatment impact. Collecting resource use and cost information in the context of a clinical trial, as was done in constructing the data set for this study, is an extremely cumbersome and costly undertaking. The equations presented here, however, permit us to reduce, and perhaps forego entirely, this step and to directly estimate the costs on the basis of the patient's early results. Although the equations do not directly examine the economic impact of treatments that influence the severity of impairment and death, they do afford a means of estimating the costs to be expected during the first 12 weeks after stroke onset. When using the regression equations for this purpose, one should recognize that treatment may not only change the severity of impairment early on but may change how the predictors interact—in effect altering the equations themselves.

Caution is warranted when making predictions in extreme strata of the population, although the underlying data for our models reflect very wide ranges of all the determinants.

Our findings suggest that prospective reimbursement systems for stroke rehabilitation should factor in the patients' functional status and severity of the neurological impairment to provide fair reimbursement for the care of those who benefit from acute inpatient rehabilitation. For the individual clinician, the model provides a tool for projecting the resource needs for stroke care and rehabilitation in the short to medium term, given the distinct characteristics of a patient population. In this respect, it should be noted that although the model was shown to be very reliable at predicting the mean cost for specific populations, substantial variations at an individual patient level are to be anticipated. This is inherent to predictive models of this sort and by no means undermines its validity or usefulness.

The correct use of regression equations when making cost predictions at a population level is important. It is common practice in the medical literature to directly enter the population's mean value (continuous variables) or the prevalence (dichotomous variables) in the equation. Results derived in this way, however, do not properly reflect the clustering of factors in the population and hence are liable to give incorrect outcomes. For example, patients with severe neurological impairment also tend to be older and to have a lower Barthel Index—a clustering that the respective means do not take into account. The total cost should be derived on the basis of the factor profiles that actually occur and their frequency, instead. Despite the increased effort involved, this is the only way to accurately predict the cost for the target population.[20]

The broad geographical scope of the trials, one of the strengths of this study, also presented some major challenges. It is well known that the absolute cost of medical care varies across countries and even regions[21,22] as the result of differences in treatment practices, availability of care facilities, cultural environment, and cost structure. One could therefore legitimately question whether the absolute costs obtained with the use of a predictive model derived from two international studies apply to any specific setting outside, or even within the range of countries that participated in the trials. To help address this issue, the absolute cost equations can be converted to a relative cost model—relative costs are expected to be more stable across countries than absolute costs[23]—whereby the cost for each patient type (the dependent variable) is expressed as a multiple of the local cost of a predefined reference patient group. Because the local cost for the reference group can be calculated with up-to-date data, this relative application helps insulate the equations from both local factors and time. More details on this approach can be obtained directly from the authors.

The analyses presented here provide a tool for predicting the management costs according to factors that are readily obtained in the first few days after a stroke. They also may be used to estimate the cost implications of changes in neurological impairment induced by treatment or otherwise. It is hoped that this model will be useful at various levels in the debate about the appropriate allocation of limited healthcare resources—an ongoing challenge against a background of ever-increasing tension between demand and supply.

Appendix

Members of the Stroke Economic Analysis Group are John Cai, PhD (currently affiliated with Massachusetts Division of Health Care Financing and Policy), J. Jaime Caro, MDCM (Chairman), Krista F. Huybrechts, MS, Wendy S. Klittich, BS, Judith A. O'Brien, RN, and Gabriel Raggio, ScD, Caro Research, Concord, Mass; Ingrid Caro, MEd, Caro Research, Montreal, Quebec, Canada; and Inge Duchesne, MS, Janssen Research Foundation, Beerse, Belgium.

Acknowledgments

We are grateful to the Janssen Research Foundation (Belgium) for providing the data. The authors wish to thank Prof Willard G. Manning and Christian Kronborg Andersen for insightful comments and exchanges.

References

1. Holloway RG, Benesch CG, Rahilly CR, Courtright CE. A systematic review of cost-effectiveness research of stroke evaluation and treatment. *Stroke*. 1999;30:1340–1349.
2. Elixhauser A, ed. Health care cost-benefit and cost-effectiveness analysis (CBA/CEA) from 1991 to 1996: an updated bibliography: appendix C. *Med Care*. 1998;36:MS1–MS9.
3. Grotta J, for the US and Canadian Lubeluzole Ischemic Stroke Study Group. Lubeluzole treatment of acute ischemic stroke. *Stroke*. 1997;28:2338–2346.
4. Diener HC, for the European and Australian Lubeluzole Ischemic Stroke Study Group. Multinational randomized trial of lubeluzole in acute ischemic stroke. *Cerebrovasc Dis*. 1998;8:172–181.

5. Caro JJ, Huybrechts KF, Duchesne I, for the Stroke Economic Analysis Group. Management Patterns and Costs of Acute Ischemic Stroke: an international study. *Stroke*. 2000;31:582–590.
6. Netten A, Dennett J. *Unit Costs of Health and Social Care*. Canterbury, UK: University of Kent, Personal Social Services Research Unit; 1996.
7. National Health Service. *Scottish Health Service Costs*. Edinburgh, Scotland: Information and Statistics Division, National Health Service; 1995.
8. Willke RJ, Glick HA, Polsky D, Schulman K. Estimating country-specific cost-effectiveness from multinational clinical trials. *Health Econ*. 1998; 7:481–493.
9. Adams HP Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh E, and TOAST Investigators. Classification of subtype of acute ischemic stroke: definitions for use in a multicenter clinical trial. *Stroke*. 1993;24:35–41.
10. Strand T, Asplund K, Eriksson S, Hägg E, Lithner F, Wester PO. Stroke unit care: who benefits? Comparisons with general medical care in relation to prognostic indicators on admission. *Stroke*. 1986;17:377-381.
11. Stroke Units Trialists' Collaboration. Collaborative systematic review of the randomized trials of organized inpatient (stroke unit) care after stroke. *BMJ*. 1997;314:1151–1159.
12. Jørgenson HS, Nakayama H, Raaschou HO, Olsen TS. Acute stroke care and rehabilitation: an analysis of the direct cost and its clinical and social determinants: the Copenhagen Stroke Study. *Stroke*. 1997;28:1138–1141.
13. Kleinbaum DG, Kupper LL, Muller KE, eds. *Applied Regression Analysis and Other Multivariable Methods*. Boston, Mass: PWS-KENT Publishing Co; 1988.
14. Rutten-Van Mölken MPMH, Van Doorslaer EKA, Van Vliet RCJA. Statistical analysis of cost outcomes in a randomized controlled clinical trial. *Health Econ*. 1994;3:333–345.
15. Duan N. Smearing estimate: a nonparametric retransformation method. *J Am Stat Assoc*. 1983;78:605–610.
16. Manning WG. The logged dependent variable, heteroscedasticity, and the retransformation problem. *J Health Econ*. 1998;17:283–295.
17. Harvey RL, Roth EJ, Heinemann AW, Lovell LL, McGuire JR, Diaz S. Stroke rehabilitation: clinical predictors of resource utilization. *Arch Phys Med Rehabil*. 1998;79:1349–1355.
18. Galski T, Bruno RL, Zorowitz R, Walker J. Predicting length of stay, functional outcome, and aftercare in the rehabilitation of stroke patients: the dominant role of higher-order cognition. *Stroke*. 1993;24:1794–1800.
19. Wade DT, Collin C. The Barthel ADL Index: a standard measure of physical disability? *Int Disabil Studies*. 1988;10:64–67.
20. Caro J, Huybrechts K, Klittich W. *Pitfalls When Using Regression Equations to Estimate Treatment Impact in Actual Practice*. Workshop at annual conference of the International Society for Pharmacoeconomics and Outcomes Research; May 1999; Arlington, Va.
21. Welch WP, Miller ME, Welch HG, Fisher ES, Wennberg JE. Geographic variation in expenditures for physician's services in the United States. *N Engl J Med*. 1993;328:621–627.
22. Rouleau JL, Moyé LA, Pfeffer MA, Arnold JMO, Bernstein V, Cuddy TE, Dagenais GR, Geltman EM, Goldman S, Gordon D, Hamm P, Klein M, Lamas GA, McCans J, McEwan P, Menapace FJ, Parker JO, Sestier F, Sussex B, Braunwald E, for the SAVE Investigators. A comparison of management patterns after myocardial infarction in Canada and the United States. *N Engl J Med*. 1993;328:779–784.
23. Ramsey SD, Newton K, Blough D, McCulloch DK, Sandhu N, Wagner EH. Patient-level estimates of the cost of complications in diabetes in a managed-care population. *Pharmacoeconomics*. 1999;16:285–295.
24. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York, NY: Chapman and Hall; 1993.

Predicting Treatment Costs After Acute Ischemic Stroke on the Basis of Patient Characteristics at Presentation and Early Dysfunction

J. Jaime Caro, Krista F. Huybrechts and Heather E. Kelley
for the Stroke Economic Analysis Group

Stroke. 2001;32:100-106
doi: 10.1161/01.STR.32.1.100
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 0039-2499. Online ISSN: 1524-4628