Inter-individual Variability in the Capacity for Motor Recovery After Ischemic Stroke

Shyam Prabhakaran, Eric Zarahn, Claire Riley, Allison Speizer, Ji Y. Chong, Ronald M. Lazar, Randolph S. Marshall and John W. Krakauer

Neurorehabil Neural Repair 2008; 22; 64 originally published online Aug 8, 2007;
DOI: 10.1177/1545968307305302

Inter-individual Variability in the Capacity for Motor Recovery After Ischemic Stroke

Shyam Prabhakaran, MD, Eric Zarahn, PhD, Claire Riley, MD, Allison Speizer, MS, Ji Y. Chong, MD, Ronald M. Lazar, PhD, Randolph S. Marshall, MD, MS, and John W. Krakauer, MD

Background. Motor recovery after stroke is predicted only moderately by clinical variables, implying that there is still a substantial amount of unexplained, biologically meaningful variability in recovery. Regression diagnostics can indicate whether this is associated simply with Gaussian error or instead with multiple subpopulations that vary in their relationships to the clinical variables. *Objective.* To perform regression diagnostics on a linear model for recovery versus clinical predictors. *Methods.* Forty-one patients with ischemic stroke were studied. Impairment was assessed using the upper extremity Fugl-Meyer Motor Score. Motor recovery was defined as the change in the upper extremity Fugl-Meyer Motor Score from 24 to 72 hours after stroke to 3 or 6 months later. The clinical predictors in the model were age, gender, infarct location (subcortical vs cortical), diffusion weighted imaging infarct volume, time to reassessment, and acute upper extremity Fugl-Meyer Motor Score. Regression diagnostics included a Kolmogorov-Smirnov test for Gaussian errors and a test for outliers using Studentized deleted residuals. *Results.* In the random sample, clinical variables explained only 47% of the variance in recovery. Among the patients with the most severe initial impairment, there was a set of regression outliers who recovered very poorly. With the outliers removed, explained variance in recovery increased to 89%, and recovery was well approximated by a proportional relationship with initial impairment (recovery $\cong 0.70 \times$ initial impairment). *Conclusions.* Clinical variables only moderately predict motor recovery. Regression diagnostics demonstrated the existence of a subpopulation of outliers with severe initial impairment who show little recovery. When these outliers were removed, clinical variables were good predictors of recovery among the remaining patients, showing a tight proportional relationship to initial impairment.

Key Words: *Predictors of Recovery—Hemiparesis—Impairment—Fugl-Meyer Scale.*

From The Neurological Institute, Columbia University, Stroke and Critical Care Division, New York, New York.

Address correspondence to John W. Krakauer, MD, The Neurological Institute, Columbia University, Stroke and Critical Care Division, 710 W 168th Street, Room 640, New York, NY 10032. E-mail: jwk18@columbia.edu.

Prabhakaran S, Zarahn E, Riley C, et al. Inter-individual variability in the capacity for motor recovery after ischemic stroke. *Neurorehabil Neural Repair* 2008;22:64-71.

DOI: 10.1177/1545968307305302

Publications about stroke recovery rarely make explicit mention of unaccounted for variability in stroke recovery, and even when they do, it is often without an available citation. This may be because variability in stroke recovery is considered a clinically obvious fact and because previous studies of stroke recovery using regression models have emphasized the amount of variance successfully explained by the predictors rather than considering the biological significance of residual unexplained variance. Several observational studies have demonstrated that initial severity of hemiparesis, measured with either disability or impairment scales, is the best predictor of recovery. [1-5] Other predictors used in regression models have included age, demographics, nonmotor clinical features, infarct side and location, and stroke subtype.[1-4,6] Heretofore, recovery has usually been defined with a functional outcome scale. However, most relevant to the current report are those few studies that have defined recovery with an impairment scale such as the Fugl-Meyer Motor Score. These studies have shown that prediction of recovery by initial severity (measured within 24 hours of stroke) is only moderate, with an estimated R^2 of 30% to 50%.[2,6]

The reliability of a measurement is the ratio of true inter-individual variance to total variance and so varies from 0 to 1. The closer the reliability is to 1, the greater the extent to which variability in the measurement reflects true inter-individual variability. The Fugl-Meyer Motor Score has been shown to have a reliability close to one.[7] Therefore, the substantial amount of variability in recovery unexplained by initial severity is not simply measurement error but biologically meaningful variation. Regression diagnostics have not been reported, thus leaving unaddressed the issue of model validity (in particular that errors are identically distributed Gaussian random variables). This is of practical importance because we need to understand better the expected improvement for individual patients. Variability in recovery may be concentrated in certain subgroups and not reflect uniform uncertainty across all patients. To examine regression model validity, we performed basic

regression diagnostics on prediction of motor recovery by clinical variables.

In this article, we define our recovery measure Δ_{FM} as the Fugl-Meyer Motor upper extremity score (FM-UE) measured at approximately 3 and 6 months postinfarct minus the FM-UE at 24 to 72 hours postinfarct. Thus, positive values of Δ_{FM} indicate recovery. We used an impairment scale, FM-UE, because we believe that it is reduction of impairment rather than of disability (the latter often being mediated through compensatory mechanisms)[7,8] that reflects true biological recovery. Furthermore, Δ_{FM} rather than final FM-UE was chosen because a recovery process fundamentally implies a change in state rather than just an endpoint. By analogy, a skin laceration and a graze may both fully heal, but the laceration engages the biological mechanisms of healing more than the graze.

SUBJECTS AND METHODS

The Performance and Recovery in Stroke (PARIS) registry is a prospective natural history study of neurologic deficits following first-time stroke, using measures of impairment and structural brain imaging. Between May 2002 and May 2006, eligible patients were screened from the Columbia University Medical Center inpatient stroke service after a diagnosis of ischemic stroke was confirmed. Those who were able to consent and willing to participate in the study were enrolled within 72 hours of stroke onset. From the overall cohort of 87 patients enrolled to date, there were 41 subjects with an initial FM-UE <66 (PARIS motor cohort).

Through medical chart review and patient interview, we collected baseline information on age, sociodemographic history, medication history, and vascular risk factors (current smoking, tobacco use, atrial fibrillation, hypertension, diabetes mellitus, hypercholesterolemia, and coronary artery disease). Initial examination, performed between 24 and 72 hours of stroke onset, included a detailed neurologic examination, the National Institutes of Health Stroke Scale Score, manual motor testing using the Medical Research Council rating, and the Fugl-Meyer Scale.

Radiographic data on infarct topography, volume, and arterial territory were also obtained between 24 and 72 hours of stroke onset. Lesion volume was estimated in cm^3: lesion volume = [product of maximal perpendicular diameters of the diffusion weighted imaging (DWI) lesion in cm] × [number of 0.5-cm slices]/2, a reasonably reliable method compared with automated methods.[9,10] Lesion location was dichotomized into subcortical only versus any cortical involvement. The diagnostic workup to determine stroke etiology was made by the treating physician and was not standardized. Using the Trial of ORG 10172 in Acute Stroke Treatment criteria based on the history, clinical, and diagnostic evaluation, we assigned stroke subtype classification for each stroke.[11] Each patient in this study had physical and occupational therapy during his or her hospitalization.

All patients were expected to have follow-ups at 3 and 6 months after their stroke, when neurological examinations, manual motor testing, and Fugl-Meyer Scale would be performed. The time to reassessment

Table 1. Patient Demographics and Clinical Characteristics (N = 41)

Age, y	
Mean (SD)	59.7 (10.1)
Median	59.0
Male, n (%)	23 (56.1)
Race, n (%)	
White	12 (29.3)
Black	15 (36.6)
Hispanic	12 (29.3)
Other	2 (4.8)
Hypertension, n (%)	31 (75.6)
Diabetes, n (%)	17 (41.5)
Dyslipidemia, n (%)	14 (34.1)
Atrial fibrillation, n (%)	6 (14.6)
Antithrombotic use, n (%)	13 (31.7)
Statin use, n (%)	12 (29.3)
Smoking, n (%)	10 (24.4)
Stroke subtype, n (%)	
Small vessel	17 (41.5)
Large vessel	7 (17.1)
Cardioembolic	7 (17.1)
Cryptogenic	10 (24.4)
Infarct location, n (%)	
Small (<2 cm^3) deep subcortical only	18 (43.9)
Superficial or large (>2 cm^3) deep subcortical only	9 (22.0)
Cortical ± immediate subcortical	14 (34.2)
FM-UE	
Mean (SD)	32.5 (25.3)
Median	31
Baseline NIHSS score	
Mean (SD)	5.9 (4.1)
Median	5.0
DWI volume (cm^3)	
Mean (SD)	3.1 (2.8)
Median	1.8
Maximum FM-UE	
Mean (SD)	50.9 (19.5)
Median	61
Follow-up time, y	
Mean (SD)	0.33 (0.02)
Median	0.27

FM-UE = Fugl-Meyer Motor Score—upper extremity; NIHSS = National Institutes of Health Stroke Scale; DWI = diffusion weighted imaging.

Table 2. Estimated Regression Coefficients for Δ_{FM} as the Dependent Variable (N = 41)

	Coefficient Estimate	Standard Error	$t(32)$	2-Tailed P
y-intercept	35.06	25.03	1.40	1.7×10^{-1}
Acute FM-UE	–0.40	0.093	–4.34	1.3×10^{-4}
Lesion location (cortical = 1; subcortical = 0)	–3.29	8.39	–0.39	7.0×10^{-1}
Cortical lesion volume	0.13	1.24	0.11	9.2×10^{-1}
Subcortical lesion volume[a]	–3.46	1.34	–2.59	1.4×10^{-2}
Age	–0.21	0.37	–0.57	5.7×10^{-1}
Gender (male = 1; female = 0)	17.15	29.75	0.58	5.7×10^{-1}
Age × gender	–0.24	0.49	–0.49	6.3×10^{-1}
Time to reassessment	40.64	20.10	2.02	5.0×10^{-2}

FM-UE = Fugl-Meyer Motor Score—upper extremity.

(which was included as a predictor) varied by 2 to 4 weeks around the desired date of follow-up. In addition, not every patient had both follow-ups. We used the 6-month data in those patients who had both follow-ups. The maximum possible FM-UE score was 66. As stated earlier, Δ_{FM} was the chosen measure of motor recovery.

Linear regression (N = 41) was used to model Δ_{FM} with currently known or plausible clinical predictors: age, gender, age × gender, infarct location (subcortical vs cortical), DWI infarct volume, infarct location × infarct volume, time to reassessment, and acute FM-UE. The false positive rate was controlled at $\alpha = .05$ per predictor.

An independent estimate of the measurement error variance of Δ_{FM} was obtained from the literature and equaled 8.4 (N = 22).[12] The ratio of a sample variance of Δ_{FM} to this measurement error variance yields an F statistic that can be used to test whether reliability > 0. Reliability was directly computed from the F statistic as $1 - 1/F$.

Regression diagnostics included a Kolmogorov-Smirnov test for normality of Studentized residuals, using Studentized deleted residuals to test for outliers and plotting Studentized residuals versus the predicted value of the model.

RESULTS

The patients' (N = 41) baseline characteristics are summarized in Table 1. The average (±SD) Δ_{FM} was 18.4 ± 18.5 (range 0-61). Δ_{FM} was significantly positive ($t_{40} = 6.4$, $P < 10^{-6}$). The estimated reliability of Δ_{FM} was 0.98, which was significantly greater than zero ($F_{40,21} = 40.6$, $P < 10^{-6}$).

Δ_{FM} was then regressed against the predictor variables stated in the Methods (Table 2). The overall model was significant ($F_{8,32} = 5.3$, $P = 2.6 \times 10^{-4}$). Only acute FM-UE and subcortical lesion volume were individually significant. The statistical results for each predictor are shown in Table 2. The adjusted R^2 for the model was 0.47 (nonadjusted $R^2 = 0.57$). The mean squared residual for this model was 182.6 with 32 df, corresponding to a reliability estimate of 0.95, which was significantly greater than zero ($F_{32,21} = 21.7$, $P < 10^{-6}$). Thus, an estimated 95% of the unexplained variance in this regression model is attributable to true inter-individual variability and 5% is attributable to measurement error.

A Kolmogorov-Smirnov test performed on the Studentized residuals showed a trend toward violation of normality (Kolmogorov-Smirnov statistic = 0.135, $df = 41$, $P = .06$). A plot of Studentized residuals versus predicted Δ_{FM} clearly indicated heteroscedasticity (Figure 1). A plot of predicted versus observed Δ_{FM} suggested a poorer model fit at smaller values of Δ_{FM} (Figure 2); in particular, the variance of the residuals seemed substantially larger at values of $\Delta_{FM} < 18$. This was confirmed by an F test that compared the variances of the Studentized residuals between $\Delta_{FM} < 18$ and $\Delta_{FM} \geq 18$ ($F_{24,15} = 3.42$, $P = 8.3 \times 10^{-3}$). Inspection of Figure 2 did not suggest simply a greater error variance at lower values of Δ_{FM} but rather that there were several outliers lying above a regression line clearly evident for $\Delta_{FM} \geq 18$. A standard test for outliers (Studentized deleted residuals) detected only 1 outlier (a patient with acute FM-UE = 4 and $\Delta_{FM} = 8$), but this test is not sensitive to the presence of multiple outliers. To formalize the impression that there was a subgroup of outliers to an otherwise tight regression line, we estimated linear model parameters based only on subjects whose $\Delta_{FM} \geq 18$ (these subjects all had acute FM-UE ≤ 45; N = 17). Unsurprisingly, the overall fit from this model to the $\Delta_{FM} \geq 18$ data was excellent (nonadjusted $R^2 = 0.93$, adjusted $R^2 = 0.85$; Figure 3). From these parameter estimates, we then computed Studentized residuals for $\Delta_{FM} < 18$ (N = 24); the proportion of these residuals that were negative (ie, corresponding to observed values that were less than their predicted values) was significantly greater than

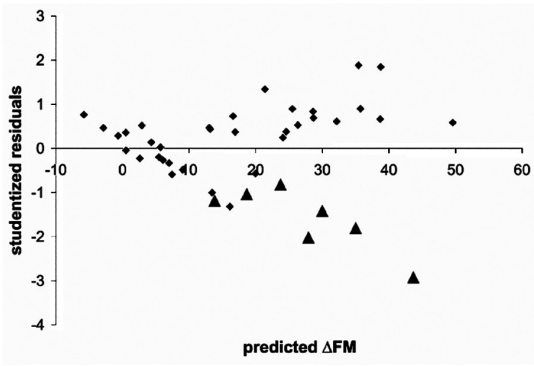

Figure 1. Studentized residuals versus standardized regression predicted values for the regression model of Table 2. Heteroscedasticity is supported. The triangles indicate the outliers identified subsequently.

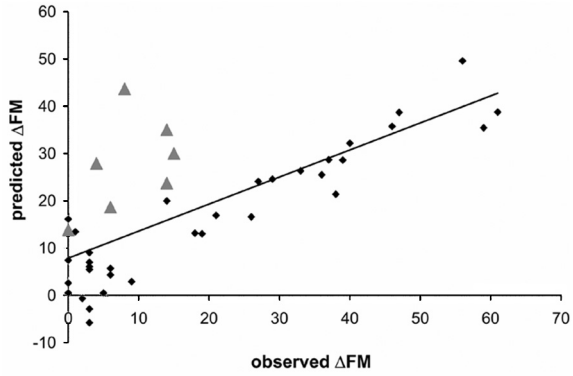

Figure 2. Predicted versus observed Δ_{FM} for the regression model of Table 2. The line is the least-squares fit. The triangles indicate the outliers identified subsequently.

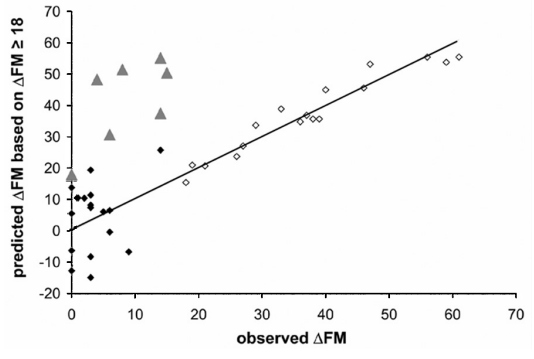

Figure 3. The open diamonds are the data for which $\Delta_{FM} \geq 18$. The same regression model as in Table 2 was fit to these data and then used to compute fits for $\Delta_{FM} < 18$ (solid diamonds and triangles). The identity line is plotted to illustrate the magnitude and sign of the residuals: the sign of a residual depends on whether it is above (negative, by convention) or below (positive) this line. The triangles indicate the outliers identified using this regression equation.

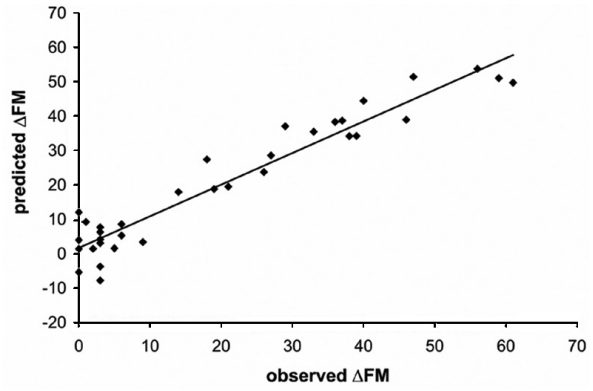

Figure 4. Predicted versus observed Δ_{FM} for the regression model of Table 3 (excluding the outliers). The line is the least-squares fit.

50% (2-tailed $P = .02$ by binomial test). Also, the largest negative Studentized residual was –7.25, whereas the largest positive Studentized residual was only 1.90. These diagnostics are consistent with the existence of a subgroup of patients who do not follow the main regression equation for recovery and instead show lower recovery than predicted.

We then identified members of this subgroup based on their Studentized residuals computed from the regression equation estimated from $\Delta_{FM} \geq 18$: all Studentized residuals whose 2-tailed P values were less than .05 (corresponding to $|t(8)| > 2.31$) were classified as outliers. This procedure yielded 7 such cases (Figure 3). The outlier subjects were notable for a tendency to have low values of acute FM-UE (mean = 5.9 ± 9.2 compared with 38.0 ± 24.0 for the nonoutlier subjects), which was the largest determinant of predicted Δ_{FM} (Table 2). Thus, subjects with low acute FM-UE tended to have the greatest inter-individual variability, because some of them recovered a great deal whereas others (the outliers) recovered very little.

The overall model was highly significant ($F_{8,25} = 35.7$, $P < 10^{-6}$) when the regression was recomputed without the 7 outliers (Figure 4). There were significant individual predictor effects for acute FM-UE, subcortical lesion volume, age, and y-intercept (Table 3). The adjusted R^2 for this model = 0.89 (nonadjusted $R^2 = 0.92$), a large improvement in proportion of explained variance compared with when the outliers were included (Figure 3). A Kolmogorov-Smirnov test performed on the Studentized residuals to test for non-normality was not significant (Kolmogorov-Smirnov statistic = 0.08, $df = 34$, $P = .20$), and neither skewness nor kurtosis was significantly different from 0.

Even with the much greater explanatory power of this regression model (mean squared residual = 40.8 with 25

Table 3. Estimated Regression Coefficients for Δ_{FM} as the Dependent Variable Excluding 7 Outlier Cases ($N = 34$)

	Coefficient Estimate	Standard Error	$t(25)$	2-Tailed P
y-intercept[a]	72.58	13.39	5.42	1.3×10^{-5}
Acute FM-UE[a]	–0.69	0.06	–12.57	$<10^{-6}$
Lesion location (cortical = 1; subcortical = 0)	5.33	5.07	1.05	3.0×10^{-1}
Cortical lesion volume	–0.827	0.743	–1.11	2.8×10^{-1}
Subcortical lesion volume[a]	–2.97	0.73	–4.07	4.1×10^{-4}
Age[a]	–0.50	0.19	–2.57	1.7×10^{-2}
Gender (male = 1; female = 0)	–12.52	15.85	–0.79	4.4×10^{-1}
Age × gender	0.24	0.26	0.92	3.7×10^{-1}
Time to reassessment	19.48	11.21	1.74	9.5×10^{-2}

FM-UE = Fugl-Meyer Motor Score—upper extremity.
a. 2-tailed $P < .05$.

Table 4. Estimated Regression Coefficients for Δ_{FM} as the Dependent Variable Excluding 7 Outlier Cases (N = 34) and After Paring Model Stepwise Until All Remaining Predictors Were Significant at $\alpha = .05$

	Coefficient Estimate	Standard Error	$t(29)$	2-tailed P
y-intercept	63.23	7.49	8.44	$<10^{-6}$
Acute FM-UE	–0.70	0.05	–14.76	$<10^{-6}$
Subcortical lesion volume	–2.94	0.60	–4.88	3.5×10^{-5}
Age	–0.32	0.11	–2.78	9.4×10^{-3}
Time to reassessment	19.55	9.57	2.04	5.0×10^{-2}

FM-UE = Fugl-Meyer Motor Score—upper extremity.

df), the reliability estimate of the residuals was still high (reliability estimate of 0.79) and significantly greater than zero ($F_{25,21} = 4.9$, $P = 2.5 \times 10^{-4}$). Thus, most of the variability in Δ_{FM} remained attributable to true inter-individual variability. These results indicate that outliers to the regression equation are the cause of the bulk of the unexplained variance in recovery (mean squared residual = 182.6 with outliers vs 40.8 without outliers).

To obtain a more parsimonious regression model (again excluding the same 7 outliers), predictors were backward eliminated stepwise until all remaining predictors were significant at $\alpha = .05$ or until the adjusted R^2 decreased. The final selected model had eliminated lesion location, cortical lesion volume, gender, and age-by-gender interaction and retained acute FM-UE, subcortical lesion volume, age, time to reassessment, and y-intercept (Table 4). The adjusted R^2 for this model = 0.90 (nonadjusted R^2 = 0.91). When applied to the 7 outliers, this same reduced model was not significant ($F_{4,2} = 2.6$, $P = .29$).

We defined maximal potential recovery as 66 – acute FM-UE. To obtain a prediction of the proportion of this difference that patients will recover, we evaluated the reduced model at the mean values of subcortical lesion volume (1.36 cm^3), age (59.8 years), and time to reassessment (121 days). This gave an estimated regression model of $\Delta_{FM} = (0.70) \cdot (66 - \text{acute FM-UE}) + 0.4 \approx (0.70) \cdot (\text{maximal potential recovery})$. Thus, with the exception of the outlier subpopulation, patients recover in a proportional manner, tending to realize approximately 70% of their maximal potential recovery.

DISCUSSION

We sought to formally determine the sources and structure of inter-individual variability in motor recovery after stroke. We found, replicating previous results,[7] that the reliability of the FM-UE was almost 1, implying that inter-individual differences almost entirely reflect biologically meaningful variability. Two sources were found to account for nearly all this systematic variation. One was initial impairment (other clinical variables also played a role), with which recovery in the majority of subjects exhibited a nearly proportional relationship (proportional recovery). Another source, discovered through regression diagnostics, was the existence of a subgroup of patients with high initial impairment that did not show proportional recovery, corroborating an earlier study which observed that some patients with severe initial deficits do not recover whereas others do.[13]

Except for these outliers, most patients tended to recover approximately 70% of their initial impairment

(proportional recovery). Proportionality was not simply attributable to a ceiling effect for the FM in patients with minimal initial impairment, because this relationship was identified strongly even in subjects with acute FM-UE ≤ 45 (and $\Delta_{FM} \geq 18$; compare line fits for Figures 3 and 4). A striking implication of this finding is that even though damage to the corticospinal tract (CST) is largely irreversible, there is a mechanism that nevertheless allows substantial recovery even from severe impairment. One might think that proportional recovery is a necessary consequence of any relationship between recovery and initial impairment or is an inevitable result of outlier exclusion. Neither intuition would be correct. The slope (m) plus y-intercept (b) model relating recovery (R) to initial impairment (I) is $R = mI + b$. The free parameters estimated by regression are m and b. Proportional recovery is a special case of this relationship with $b = 0$; this is a function of the data and not a foregone conclusion of the modeling procedure itself (ie, if recovery data were generated by a different process, b could certainly have taken on substantial non-zero values, in which case recovery would not have been well approximated as proportional). For the same reason, exclusion of outliers could not itself cause a proportional relationship but only improve the regression fit to the model $R = mI + b$.

A possible mechanism for the failure of outliers to recover is that these patients have less residual CST integrity after stroke compared with those patients with similar initial severity but who show proportional recovery, although because the model included DWI volume, this difference in CST integrity would have to be unrelated to DWI lesion volume. A relationship of CST integrity to prognosis is suggested by studies using diffusion-tensor imaging[14,15] and corticospinal masks,[16] which show that the degree of CST injury correlates negatively with motor recovery following stroke.[14,16]

Further support comes from studies using transcranial magnetic stimulation (TMS) with motor evoked potentials (MEPs) measurements in the first 10 days after stroke. One review reported that TMS in patients with severe acute impairment is a highly specific test of future recovery, which would imply that any severely impaired patient who does not recover will have absent MEPs in the acute stroke period. Therefore, based on this idea, it would be predicted that our outliers (all severely impaired patients who recovered poorly) would have absent MEPs. However, neither specificity nor sensitivity of TMS has been uniformly high in the literature, with great variability between laboratories.[17] For example, some studies of MEPs response very early after stroke have not been able to distinguish between good and poor recoverers.[18] To conclusively determine whether the regression outliers had absent MEPs would require a study combining both acute stroke period TMS and clinical variable regression diagnostics.

Even if it is the case that CST integrity is the final mediator of recovery, it is critical to ask why there is a delay between the detection of MEPs and recovery. It is not likely that edema and metabolic derangement, which are maximal within 72 hours,[19] are the only explanations. We suggest that another mechanism contributing to recovery is the capacity of the remaining undamaged brain to reorganize and subsequently recruit the residual CST and other descending pathways. This speculation is consistent with positron emission tomography and functional magnetic resonance imaging studies showing that both ipsilesional and contralesional areas are activated during recovery from hemiparesis.[20,21] It can be speculated that patients in whom MEPs are present in the acute stroke phase but who do not recover[22] lack this specific capacity for brain reorganization.

Although many previous studies have shown that initial stroke severity is the best predictor of subsequent functional outcome, only a few studies have used measures of impairment rather than disability. Duncan and colleagues[2] studied the relationship between initial and 6-month total FM in 54 patients. They obtained an R^2 of 0.53 when the initial FM was measured in the first 24 hours after stroke.[2] Similarly, Feys et al[6] obtained an R^2 of 0.30, although the follow-up FM was at 12 rather than 6 months.[6] Neither of these studies reported regression model diagnostics. Furthermore, although the FM-UE has been reported to have excellent reliability (>0.99),[7] the 2 cited studies defined recovery as the final FM-UE, whereas we defined recovery as Δ_{FM}. This is not trivial, because changes in performance and final performance are only moderately correlated, and dynamic recovery processes during the acute stroke period would be expected to relate to future changes in performance rather than final levels of performance per se. To our knowledge, this is the first time that the reliability of Δ_{FM} has been reported.

Our finding that 95% of variance in recovery unexplained by clinical variables is attributable to true inter-individual variability implies that there are as-yet unidentified biological processes that account for observed inter-individual differences. More precisely, because the estimated unexplained variance in recovery from the regression model was 53% of the total and the reliability of this variance was estimated at 95%, our estimate of the amount of biological variability in recovery as-yet unexplained = (53%) · (95%) = 50% of the total, observed variance in Δ_{FM} in the population. On the basis of our current results, we would argue that the main source of this unexplained variability in recovery is the

presence of the outliers to the proportional recovery relationship, which seem to only exist at the severe range of initial impairment. These outliers appear to represent a clinically meaningful subpopulation distinct from the majority of patients across the entire range of initial impairment who show proportional recovery.

Limitations to this study include the small sample size, although it is similar to other series in the literature. Because of the requirements of study consent and follow-up, the majority of our patients had isolated motor deficits or had only minimal nonmotor cortical deficits such as aphasia and neglect (ie, large cortical strokes were largely excluded). Thus, our study had a selection bias toward patients with isolated hemiparesis and relatively small lesions. Future studies should examine whether proportional recovery holds for larger lesions or is modulated by additional nonmotor deficits.[23] Another concern might be that acute DWI volume without perfusion imaging can underestimate final infarct volume, but DWI has been shown to be a reasonable surrogate when performed after the first 24 hours.[24] In addition, the reliability of manual lesion measurement, although good, is not as high as with automated methods.[9,10] Last, although follow-up times were not uniform, we adjusted for time to reassessment in the regression models.

CONCLUSIONS

The results of this study lead to several important conclusions with regard to motor recovery after stroke. First, observed inter-individual variability in stroke recovery in the studied population is almost exclusively attributable to true inter-individual (ie, biologically meaningful) variability. Second, impairment in the first 72 hours (and secondarily, subcortical lesion volume, age, and time after stroke onset) is a very good predictor of recovery when high initial impairment outliers are removed. Interestingly, recovery in the nonoutliers is proportional to initial severity (with an estimated proportionality of 0.70). Third, after we accounted for clinical variables, interindividual variability seemed to be attributable primarily to the existence of a subpopulation of patients with severe initial impairment who, unlike the rest of the population (including others with severe initial impairment), do not experience proportional recovery.

ACKNOWLEDGMENTS

The work was supported by the following grants: T32-NS07153 (SP), 5R01-HD043249 (RML), P50-NS049060 (RSM), K02-NS048099 (JWK).

REFERENCES

1. Broeks JG, Lankhorst GJ, Rumping K, Prevo AJ. The long-term outcome of arm function after stroke: results of a follow-up study. *Disabil Rehabil.* 1999;21:357-364.
2. Duncan PW, Goldstein LB, Matchar D, Divine GW, Feussner J. Measurement of motor recovery after stroke: outcome assessment and sample size requirements. *Stroke.* 1992;23:1084-1089.
3. Kwakkel G, Kollen BJ, van der Grond J, Prevo AJ. Probability of regaining dexterity in the flaccid upper limb: impact of severity of paresis and time since onset in acute stroke. *Stroke.* 2003;34:2181-2186.
4. Meldrum D, Pittock SJ, Hardiman O, Ni Dhuill C, O'Regan M. Recovery of the upper limb post ischaemic stroke and the predictive value of the Orpington Prognostic Score. *Clin Rehabil.* 2004;18:694-702.
5. Nakayama H, Jorgensen HS, Raaschou HO, Olsen TS. Recovery of upper extremity function in stroke patients: the Copenhagen Stroke Study. *Arch Phys Med Rehabil.* 1994;75:394-398.
6. Feys H, De Weerdt W, Nuyens G, van de Winckel A, Selz B, Kiekens C. Predicting motor recovery of the upper limb after stroke rehabilitation: value of a clinical examination. *Physiother Res Int.* 2000;5:1-18.
7. Gladstone DJ, Danells CJ, Black SE. The Fugl-Meyer assessment of motor recovery after stroke: a critical review of its measurement properties. *Neurorehabil Neural Repair.* 2002;16:232-240.
8. Wade DT, Langton-Hewer R, Wood VA, Skilbeck CE, Ismail HM. The hemiplegic arm after stroke: measurement and recovery. *J Neurol Neurosurg Psychiatry.* 1983;46:521-524.
9. Gomez-Marino R, Andre C, Novis SA. Volumetric determination of cerebral infarction in the acute phase using skull computed tomography without contrast: comparative study of 3 methods. *Arq Neuropsiquiatr.* 2001;59:380-383.
10. van der Worp HB, Claus SP, Bar PR, et al. Reproducibility of measurements of cerebral infarct volume on CT scans. *Stroke.* 2001;32:424-430.
11. Adams HP Jr, Bendixen BH, Kappelle LJ, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke.* 1993;24:35-41.
12. van der Lee JH, Beckerman H, Lankhorst GJ, Bouter LM. The responsiveness of the Action Research Arm test and the Fugl-Meyer Assessment scale in chronic stroke patients. *J Rehabil Med.* 2001;33:110-113.
13. Binkofski F, Seitz RJ, Hacklander T, Pawelec D, Mau J, Freund HJ. Recovery of motor functions following hemiparetic stroke: a clinical and magnetic resonance-morphometric study. *Cerebrovasc Dis.* 2001;11:273-281.
14. Lie C, Hirsch JG, Rossmanith C, Hennerici MG, Gass A. Clinicotopographical correlation of corticospinal tract stroke: a color-coded diffusion tensor imaging study. *Stroke.* 2004; 35:86-92.
15. Konishi J, Yamada K, Kizu O, et al. MR tractography for the evaluation of functional recovery from lenticulostriate infarcts. *Neurology.* 2005;64:108-113.
16. Pineiro R, Pendlebury ST, Smith S, et al. Relating MRI changes to motor deficit after ischemic stroke by segmentation of functional motor pathways. *Stroke.* 2000;31:672-679.
17. Hendricks HT, Zwarts MJ, Plat EF, van Limbeek J. Systematic review for the early prediction of motor and functional outcome after stroke by using motor-evoked potentials. *Arch Phys Med Rehabil.* 2002;83:1303-1308.

18. Binkofski F, Seitz RJ, Arnold S, Classen J, Benecke R, Freund HJ. Thalamic metabolism and corticospinal tract integrity determine motor recovery in stroke. *Ann Neurol.* 1996;39:460-470.

19. Dirnagl U, Iadecola C, Moskowitz MA. Pathobiology of ischaemic stroke: an integrated view. *Trends Neurosci.* 1999;22: 391-397.

20. Marshall RS, Perera GM, Lazar RM, Krakauer JW, Constantine RC, DeLaPaz RL. Evolution of cortical activation during recovery from corticospinal tract infarction. *Stroke.* 2000;31:656-661.

21. Nelles G, Spiekermann G, Jueptner M, et al. Reorganization of sensory and motor systems in hemiplegic stroke patients: a positron emission tomography study. *Stroke.* 1999;30:1510-1516.

22. Heald A, Bates D, Cartlidge NE, French JM, Miller S. Longitudinal study of central motor conduction time following stroke. 2. Central motor conduction measured within 72 h after stroke as a predictor of functional outcome at 12 months. *Brain.* 1993;116:1371-1385.

23. Schiemanck SK, Post MW, Kwakkel G, Witkamp TD, Kappelle LJ, Prevo AJ. Ischemic lesion volume correlates with long-term functional outcome and quality of life of middle cerebral artery stroke survivors. *Restor Neurol Neurosci.* 2005;23:257-263.

24. Lu M, Mitsias PD, Ewing JR, et al. Predicting final infarct size using acute and subacute multiparametric MRI measurements in patients with ischemic stroke. *J Magn Reson Imaging.* 2005; 21:495-502.