Prediction of Motor Recovery Using Initial Impairment and fMRI 48 h Poststroke

Eric Zarahn[1], Leeor Alon[2], Sophia L. Ryan[2], Ronald M. Lazar[2], Magnus-Sebastian Vry[3], Cornelius Weiller[3], Randolph S. Marshall[2] and John W. Krakauer[4]

[1]The Neurological Institute, New York, NY 10032, USA, [2]Department of Neurology, Columbia University College of Physicians and Surgeons, New York, NY 10032, USA, [3]Department of Neurology, University of Freiburg and [4]Department of Neurology, Johns Hopkins University, Baltimore, MD 21287, USA

Address correspondence to Eric Zarahn. Email: eric.zarahn@gmail.com.

There is substantial interpatient variation in recovery from upper limb impairment after stroke in patients with severe initial impairment. Defining recovery as a change in the upper limb Fugl-Meyer score (ΔFM), we predicted ΔFM with its conditional expectation (i.e., posterior mean) given upper limb Fugl-Meyer initial impairment (FM_{ii}) and a putative functional magnetic resonance imaging (fMRI) recovery measure. Patients with first time, ischemic stroke were imaged at 2.5 ± 2.2 days poststroke with 1.5-T fMRI during a hand closure task alternating with rest (fundamental frequency = 0.025 Hz, scan duration = 172 s). Confirming a previous finding, we observed that the prediction of ΔFM by FM_{ii} alone is good in patients with nonsevere initial hemiparesis but is not good in patients with severe initial hemiparesis (96% and 16% of the total sum of squares of ΔFM explained, respectively). In patients with severe initial hemiparesis, prediction of ΔFM by the combination of FM_{ii} and the putative fMRI recovery measure nonsignificantly increased predictive explanation from 16% to 47% of the total sum of squares of ΔFM explained. The implications of this preliminary negative result are discussed.

Keywords: fMRI, prediction, recovery, stroke

Introduction

Stroke is the leading cause of long-term disability among adults, and hemiparesis is the most common impairment after stroke (Wolfe 2000; Krakauer 2005). Recovery from hemiparesis can be considered from the perspectives of recovery of function (i.e., regaining the ability to perform a given task, but not necessarily through the same pattern of muscle activation as used to perform it prestroke) and recovery from impairment (i.e., regaining the ability to perform a given task through the same pattern of muscle activation as used to perform it prestroke). We are concerned here with prediction of recovery from impairment (Fugl-Meyer et al. 1975) rather than of recovery of function. This is because though recovery of function is of great socioeconomic importance, it is influenced by compensatory strategies (Lyle 1981; van der Lee et al. 1999; Kwakkel et al. 2004).

A substantial proportion of initial impairment is recovered in patients with first time, nonsevere hemiparesis by 3 months poststroke. Using the Fugl-Meyer scale, for example, we previously observed that a change in measure of recovery (measured as the change from initial impairment to impairment at 3 months poststroke) in such patients was well described as 0.70·initial impairment (Prabhakaran et al. 2008) ("proportional recovery"). Also, recovery at this time point tends to be near (say within 10% of) asymptotic recovery in this patient subpopulation (Duncan et al. 1992, 1994; Nakayama et al. 1994; Jorgensen et al. 1995, 1999; Kwakkel et al. 2004, 2006; Swayne

et al. 2008; Verheyden et al. 2008; van Kuijk et al. 2009). The degree of regularity of the magnitude and time course of this recovery and its seeming robustness to rehabilitation dose (reviewed by Kwakkel et al. 2004) have suggested that it is mediated by a common, "spontaneous" mechanism.

In contrast to patients with nonsevere initial hemiparesis, ~50% of stroke survivors that have severe initial hemiparesis remain so in the chronic state (Nakayama et al. 1994; Jorgensen et al. 1999; Hendricks et al. 2002; Kwakkel et al. 2003; Prabhakaran et al. 2008; van Kuijk et al. 2009). Motor-evoked potentials (MEPs) at the abductor digiti minimi using transcranial magnetic stimulation (TMS) 1 week poststroke in patients with severe initial impairment have a positive predictive value of ~0.95 for recovery of some criterion dexterity at 3–6 months poststroke but tend to have less impressive negative predictive value (Hendricks, Pasman, Merx, et al. 2003; Hendricks, Pasman, van Limbeek, et al. 2003; Swayne et al. 2008; van Kuijk et al. 2009). This might 1) be an artifact of MEP threshold choice, 2) indicate that MEPs assayed within 1 week poststroke inherently cannot detect all usable, residual corticospinal (CST) connections early poststroke (perhaps due to diaschisis corticospinalis; Kwakkel et al. 2004), or 3) indicate that recovery from hand impairment can be effected through non-CST, cortical reorganization mechanisms. Furthermore, the positive predictive value of TMS-evoked MEPs has been observed to act through a delay of a few weeks such that a patient can be MEP positive while still having substantial recovery in front of him (Swayne et al. 2008; van Kuijk et al. 2009), which suggests that cortical reorganization is sometimes also required, above and beyond CST physiological integrity, to effect or complete recovery. For these reasons, if functional neuroimaging measures cortical reorganization, then it might provide information about subsequent recovery.

To wit, a multivariate correlation has been reported between early poststroke, task-related functional magnetic resonance imaging (fMRI) activation and subsequent recovery in a sample evincing a wide range of initial impairment (Marshall et al. 2009). However, the magnitude of this correlation does not directly (or at least simply) imply the accuracy of prediction of recovery based on fMRI. Here we formally assessed prediction of the same measure of recovery that we have assessed previously (Prabhakaran et al. 2008) from its posterior mean given the combined measurement of initial impairment and fMRI. We were particularly interested in the contribution of fMRI to prediction of recovery in patients with severe initial impairment as proportional recovery already seems to provide accurate prediction of recovery in patients with nonsevere initial impairment (Prabhakaran et al. 2008).

Among all possible functions of the random variables that are available as predictors (here, initial impairment and fMRI), the

posterior mean minimizes the expectation of the squared prediction error (SPE) of the random variable to be predicted (here, recovery) (Shao 2003). Computation of a posterior mean requires a conditional probability density of recovery given initial impairment and fMRI; the form of this density is thus a hypothesis being assessed by the quality of the prediction. Moreover, the free parameters of this density need in practice to be estimated. Part of this report involves describing the form of this density and the estimation of its parameters.

Materials and Methods

Ethics Statement
Patients signed informed consent forms as approved by either Columbia University's or University of Freiburg's Institutional Review Boards.

Patient Samples
Two nonoverlapping patient samples were used in this study: 1) an imaged patient sample for which prediction of recovery was performed and 2) a nonimaged patient sample from which parameters for the joint density of recovery, initial impairment, and fMRI were obtained (Fig. 1). Both samples comprised patients with first time, ischemic stroke with some degree of clinical hemiparesis (NIH stroke scale for the arm ≥ 1 for both the patients studied at Columbia Medical Center and those studied at the University of Freiburg).

The patients of the imaged sample ($N = 30$; age = 60.3 ± 9.9 years; 21 male, 28 right handed) were recruited as part of Columbia's Specialized Program of Translational Research in Acute Stroke, an National Institute of Neurological Disorders and Stroke-funded national network to investigate new pathophysiological, diagnostic, and clinical approaches in acute stroke. Fourteen of these 30 patients were part of a sample used in a previous report that did not concern imaging (Prabhakaran et al. 2008). fMRI results from 23 of these 30 patients were reported previously in a paper that did not examine prospective prediction of recovery(Marshall et al. 2009).

The nonimaged patient sample ($N = 64$; age = 61.6 ± 11.6 years; 38 M, 59 right handed) comprised 24 patients from the University of Freiburg and 40 patients from Columbia Medical Center. Twenty-seven of these 64 patients were part of a sample used in a previous report that did not concern imaging (Prabhakaran et al. 2008). Patients with prior symptomatic subcortical stroke or any prior cortical stroke were excluded. Also excluded were patients with seizure at stroke onset, aphasia, neglect, any other cognitive impairment that precluded training on the fMRI task, or any contraindication to MRI. At Columbia Medical Center, aphasia was assessed with the Western Aphasia Battery (spontaneous speech, repetition, naming, and comprehension); neglect was assessed with line bisection, letter cancellation, and line judgment; and apraxia was assessed by pantomiming scissors. At the University of Freiburg, aphasia was assessed with part 9 of the NIH stroke scale and neglect was assessed with part 11 of the NIH stroke scale; apraxia was not assessed at the University of Freiburg.

fMRI Data Acquisition
Patients underwent gradient echo echoplanar fMRI (General Electric 1.5 T, 64 × 64 matrix, field of view = 19 cm, 21 slices, slice thickness/skip = 4.5 mm/0 mm, TR = 4000 ms, TE = 52 ms, flip angle = 60°) while performing the repetitive hand closure task described below. One session (43 volumes ó 2 min 52 s) was performed per hand. fMRI data preprocessing was performed as described previously (Marshall et al. 2009).

Motor Task Used during fMRI
The motor task comprised alternations of 20-s epochs for which patients had been instructed to attempt hand closure (the first such epoch beginning 12 s after initiation of scanning) with 20-s rest epochs. Four cycles were performed per hand. The instruction for hand closure blocks was to close the hand gently from a resting position to a fist in synchrony with a 1-Hz metronome click that was played continuously (during both attempted hand closure and rest epochs) via MRI-compatible headphones in the scanner. Auditory "start" and "stop" commands were given via the headphones at the beginning and end of each 20-s attempted hand closure block. Separate runs were performed for the affected hand and unaffected hand (only the data corresponding to use of the affected hand were used in the current analysis). For patients with complete hand plegia, instructions were to "do exactly what you did with your good hand." That instructions were varied with initial impairment was understood to be an experimental design error
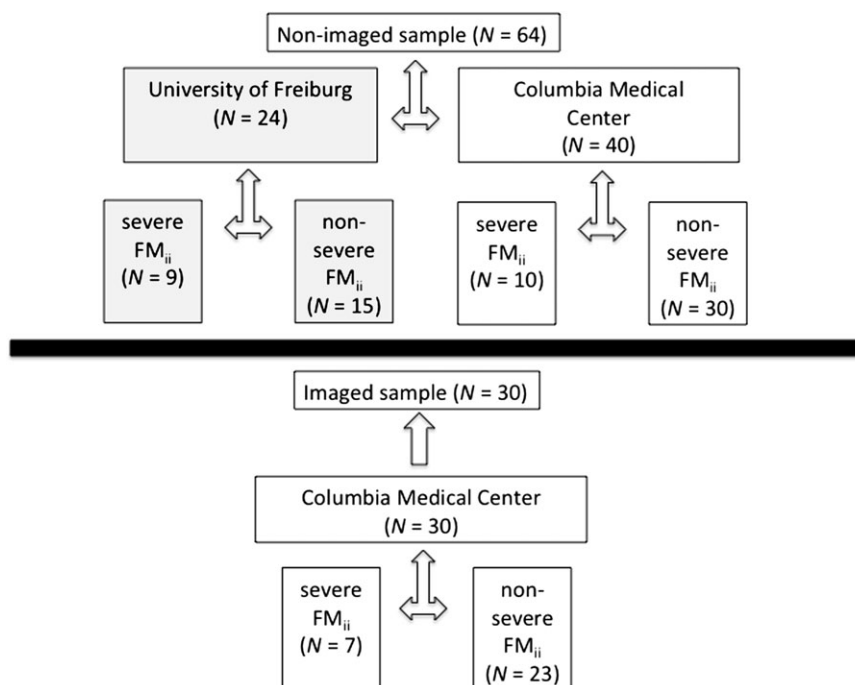

Figure 1. Schematic of patient sample composition.

in hindsight. Nevertheless, we could not conceive of any mechanism for how any confound of scanning design with initial impairment could artifactually improve prediction of recovery over initial impairment alone. Task performance was not measured, but grip force dynamometry score obtained on the day of scanning was included as a covariate in the second-level fMRI analyses (to prevent any task-related fMRI activation linearly related to dynamometry from contributing to prediction of recovery). Patients were familiarized with the task outside of the scanner. No formal assessment of mirror movements was performed; the potential effect of mirror movements on prediction is taken up in the Discussion section.

Measure of Impairment
We used the upper limb Fugl-Meyer score (FM) (Fugl-Meyer et al. 1975) as the upper limb impairment measure. The FM is a valid (Gladstone et al. 2002; Platz et al. 2005; Woodbury et al. 2008) and highly reliable (Duncan et al. 1983; Gowland et al. 1993; Gladstone et al. 2002; Platz et al. 2005; Prabhakaran et al. 2008) measure of upper limb impairment with a maximum score of 66 (higher score corresponds to less impairment). FM was assessed both at ~2 days poststroke ($FM_{initial}$) and ~3 months poststroke ($FM_{3\ months}$); more specifically, $FM_{initial}$ (as well as grip force dynamometry score) was assessed in the imaged sample on the day of fMRI scanning (2.5 ± 2.2 days poststroke). $FM_{3\ months}$ was assessed in the imaged sample at 96 ± 17 days poststroke. $FM_{initial}$ was assessed in the nonimaged sample between 1 and 3 days poststroke (average not available). $FM_{3\ months}$ was assessed in the nonimaged sample between 81 and 176 days poststroke (108 ± 25 days poststroke).

We chose 3 months as the intended endpoint because it has been shown that recovery from impairment tends to be near asymptote ≤3 months poststroke (Duncan et al. 1992; Jorgensen et al. 1995, 1999; Kwakkel et al. 2004, 2006; Swayne et al. 2008; van Kuijk et al. 2009). Initial impairment was defined as $FM_{ii} = 66 - FM_{initial}$.

Measure of Recovery
Our recovery measure was $\Delta FM = FM_{3\ months} - FM_{initial}$. We chose ΔFM as opposed to $FM_{3\ months}$ as the recovery measure because, from a physiological perspective, the process of recovery is not reflected in the final level of patient performance per se but is instead the mechanism that takes a patient from initial to final level of performance (Kwakkel et al. 2006; Prabhakaran 2008).

Dichotomous Categorization of Stroke Severity
A previous study of variability in stroke recovery (Prabhakaran et al. 2008) suggested to us that, up to some suitably high value of FM_{ii}, ΔFM was approximated well by the proportional recovery relationship:

$$\Delta FM = \beta FM_{ii}, \qquad (1)$$

with $\beta = 0.70$. At higher (i.e., more severe) values of FM_{ii}, however, this relationship qualitatively failed, with a nontrivial proportion of patients with severe FM_{ii} showing a much smaller ΔFM than that predicted by equation (1). Here, we consider the threshold FM_{ii} that determines this dichotomy in the relationship between ΔFM and FM_{ii} as a demarcation between nonsevere versus severe FM_{ii}. To get a reasonable estimate of this demarcation to be used for subsequent modeling steps, R_j (the correlation coefficient between ΔFM and FM_{ii} computed using the lowest value of FM_{ii} up to $FM_{ii,j}$, where j is an index of the ordered FM_{ii} values) was computed using the nonimaged sample and plotted against $FM_{ii,j}$. If equation (1) (for any fixed β) held throughout the entire range of FM_{ii}, then the resulting R_j versus $FM_{ii,j}$ plot would tend to stabilize around a decelerating (but increasing) curve as $FM_{ii,j}$ increased. In contrast, if there was a dichotomy as described above in the relationship between ΔFM and FM_{ii}, then the plot, after an initial period of stabilization, would begin to appreciably decrease in the neighborhood of some $FM_{ii,j}$. The plot showed this latter pattern of behavior (Fig. 2) and subjectively suggested $FM_{ii} \geq 56$ ($\Leftrightarrow FM_{initial} \leq 10$) as a reasonable definition of severe FM_{ii}. A patient with an FM of 10 in the upper limb would have a dense hemiparesis with some proximal movement but no distal movement. Using this criterion on the nonimaged sample itself yielded 46/64 patients with nonsevere FM_{ii} and 18/64 patients with severe FM_{ii}. Though it was determined in
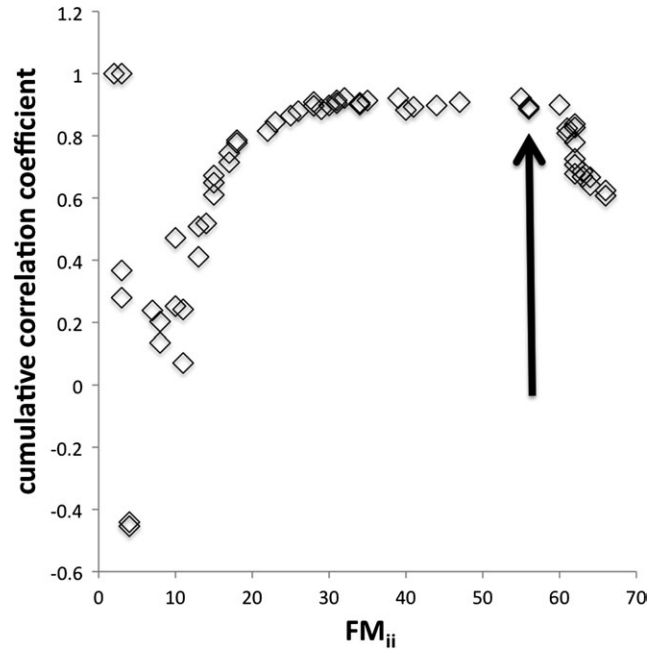

Figure 2. Cumulative correlation coefficient (see text for description) versus FM_{ii} for the nonimaged sample. The arrow marks where we specified the (inclusive) threshold (which corresponds to $FM_{ii} \geq 56$) for defining severe FM_{ii} based on where the cumulative correlation coefficient begins to decrease. We chose to place this threshold slightly before the apparent decrease so as to be less likely to misclassify severe strokes as nonsevere.

a completely different way, this criterion closely corresponds to the value of $FM_{initial}$ chosen by one other group to define severe hemiparesis (Shelton and Reding 2001), but it is more inclusive than that used by another group, which defined as severe those patients with $FM_{initial} = 0$ (van Kuijk et al. 2009).

Model for Recovery versus Initial Impairment
The conditional probability density of ΔFM given FM_{ii}, $f_{\Delta FM|FM_{ii}}$, is required for computing the posterior mean of ΔFM (see Posterior mean of ΔFM). Motivated by the previous finding described by equation (1), we assumed fundamentally a proportional recovery model,

$$\Delta FM = \beta FM_{ii} + \phi, \qquad (2)$$

with proportionality constant β and unexplained interpatient variation $\phi \sim N\left(0, \sigma_\phi^2\right)$ but with 2 flexible model features: First, it was not required that patients with nonsevere and severe FM_{ii} had the same β or σ_ϕ^2. Second, it was not required that all patients within a severity class had the same β or σ_ϕ^2. These modeling features were motivated by the data from a previous report (Prabhakaran et al. 2008), which suggested that β in patients with severe FM_{ii} took on more than one value (e.g., say 2: ~0.70 and ~0). These assumptions are formalized by stating that separate mixtures of proportional recovery models were allowed for patients with nonsevere and severe FM_{ii}; this implicitly defines $f_{\Delta FM|FM_{ii}}$. For each severity category, the number of proportional recovery models K was varied from 1 to 6 and maximum likelihood estimates of $\left(p_l, \beta_l, \sigma_{\phi,l}^2\right)_{l=1\ldots K}$ were obtained from the nonimaged sample (p_l is the weight for the lth component). The Akaike Information Criterion (AIC) (Stone 1977; Burnham and Anderson 2002) was used to select among $f_{\Delta FM|FM_{ii}}$. Details are provided in Section 3 of the Supplementary Material.

fMRI-Based Measurement of ΔFM
We obtained an fMRI-based measurement (Z_j) of ΔFM in the jth patient in the imaged sample to serve as the fMRI information in computing the posterior mean of ΔFM_j (see Posterior Mean of ΔFM). Z_j is a corrected and normalized inner product of the jth patient's fMRI "task-related activation pattern" ($\hat{a}_j$; a standard first-level statistical parametric T-map

(SPM) representing task-related activation) and a "recovery pattern" ($\hat{a}_j$; a standard second-level SPM [t] representing correlations between fMRI and ΔFM). For those unfamiliar with inner products, one can heuristically think of a regression model in which $\hat{a}_j$ acts as the dependent variable and $\hat{d}_j$ acts as the independent variable. Then Z_j can be thought of as a regression coefficient of this model that is corrected for biases. A proof that Z_j is an asymptotically unbiased estimator of ΔFM$_j$ (given the existence of some linear relationship between fMRI task-related activation and ΔFM) is provided in Section 2 of the Supplementary Material. This means that as $\hat{a}_j$ itself is estimated (see next paragraph) from more and more fMRI data sets, the difference between $E < Z_j >$ and ΔFM gets closer to zero. The formula for Z_j is provided in Section 1 of the Supplementary Material.

$\hat{d}_j$ itself is estimated from standard linear regression based on $[(\hat{a}_i, b_i, \Delta FM_i)]_{i \in A_j}$, where A$_j$ is an index set and $\mathbf{h}_i$ is an optional vector of nuisance variables from patient i (see Section 1 of Supplementary Material); here $\mathbf{h}_i$ comprised a constant term and grip force dynamometry score. To avoid artifactual contributions to the statistical relationship between Z_j and ΔFM$_j$, it is necessary that the observation $(\hat{a}_j, b_j, \Delta FM_j)$ not contribute to the computation of $\hat{d}_{-j}$. Therefore, $j \in$ A$_{-j}$ (hence the motivation for the choice of subscript for d_{-j}), which makes Z_j a cross-validatory (CV) estimator of ΔFM$_j$ (Stone 1974, 1977). This type of CV (known as either leave-one-out or N-fold CV) has the same purpose as split-half CV (Strother et al. 2002): to determine the prediction error for a dependent variable (in this case, ΔFM$_j$) when applying model parameters (in this case, $\mathbf{d}_{-j}$) estimated from an independent sample. CV is not to be confused with bootstrapping (the purpose of the latter being estimation of the variance of an estimator $\hat{\mu}$ of a model parameter μ by computing $\hat{\mu}$ repeatedly via a resampling scheme; Efron and Tibshirani 1981). The principle of CV is fundamentally related to the idea behind the AIC (Stone 1977; Burnham and Anderson 2002). If $\hat{a}_j$ were not estimated using CV, then estimation of SPE would be biased downward (i.e., it would lead to an overly optimistic assessment of prediction accuracy in future patients). The reason we used leave-one-out instead of split-half is that the former uses the maximum possible sample size for parameter estimation ($\hat{d}_{-j}$ being a multidimensional parameter) with no drawback in terms of prediction validity. Software written in MATLAB used in conjunction with SPM5 (Wellcome Department of Imaging Neuroscience) to compute Z_j as well as the fMRI data set are available from the authors upon request.

In addition to $f_{\Delta FM|FM_{ii}}$ (see Model for Recovery Versus Initial Impairment), $f_{Z|\Delta FM}$ is also required for the computation of the posterior mean of ΔFM (see Posterior mean of ΔFM). It follows from the asymptotic unbiasedness of Z that (asymptotically)

$$Z_j = \Delta FM_{jj} + \varepsilon_j, \qquad (3)$$

where ε_j is a zero mean random variable. For expedience, we assume that ε_j has a Gaussian density, which implies that asymptotically $f_{Z|\Delta FM}(Z_j|\Delta FM_j)$ is a Gaussian density with mean ΔFM$_j$ and variance σ_ε^2; the poorer the accuracy of this assumption the poorer will be the prediction ΔFM. The maximum likelihood estimator of σ_ε^2 is

$$\hat{\sigma}_\varepsilon^2 = \sum_{j=1}^{N} \frac{(\Delta FM_j - Z_j)^2}{N}. \qquad (4)$$

Posterior Mean of ΔFM

Given information about a realization of a random variable, the predictor (or, equivalently, estimator) of that random variable that minimizes the expected squared error is the expectation of that random variable with respect to the conditional density of the random variable given the information (Shao 2003). This density is called the "posterior density" in the context of Bayes' theorem, and the expectation with respect to it is called "the posterior mean." Here, we wished to predict ΔFM given the information about it contained in FM$_{ii}$ and/or Z, and so we computed the appropriate posterior means.

Assuming that Z is conditionally independent of FM$_{ii}$ given ΔFM, the posterior mean of ΔFM given FM$_{ii}$ and Z can be shown to be (see Section 3 of Supplementary Material)

$$E \langle \Delta FM | FM_{ii} = FM_{ii,j}, Z = Z_j \rangle = \frac{\int_v v f_{Z|\Delta FM}(Z_j|v) f_{\Delta FM|FM_{ii}}(v|FM_{ii,j}) \mathrm{d}v}{\int_u f_{Z|\Delta FM}(Z_j|u) f_{\Delta FM|FM_{ii}}(u|FM_{ii,j}) \mathrm{d}u}, \qquad (5)$$

where u and v are dummy variables of integration corresponding to the random variable ΔFM, which was approximated numerically. In addition to this model, 3 additional posterior means of ΔFM were computed: (conditioning on) FM$_{ii}$, Z and dichotomous stroke severity, or dichotomous stroke severity (formulae for these posterior means are provided in Supplementary eqs S3.7, 3.9, and S3.10, respectively, in Section 3 of the Supplementary Material).

For subject i of the imaged sample, the SPE of ΔFM

$$SPE_i = (\Delta FM_i - \hat{\Delta} FM_i)^2, \qquad (6)$$

where $\hat{\Delta} FM_i$, the prediction of ΔFM_i, was computed under each model as the posterior mean. The critical test was the comparison of SPE under conditioning on FM$_{ii}$ alone versus that under conditioning on FM$_{ii}$ and Z in the severe subgroup. A 1-tailed, paired t-test was used for this comparison; the rationale for using a 1-tailed test is that the alternative hypothesis is that including fMRI improves prediction (i.e., lowers SPE) over using initial impairment alone.

Results

Age (average ± standard deviation = 60.3 ± 9.9 years), time between stroke and assessment of FM$_{ii}$ (2.5 ± 2.2 days), lesion location, FM$_{ii}$ (28.9 ± 23.1), ΔFM (14.9 ± 13.8), and acute dynamometry (12.3 ± 13.3 kg) for the individual patients of the imaged sample are presented in Table 1. Patients in the imaged sample were assessed for FM$_{ii}$ and scanned with fMRI at 2.5 ± 2.2 days poststroke for 2 minutes 52 seconds per hand during the same MRI session as their clinical exams. Patients had been instructed just prior to scanning to attempt hand closure (for a specified hand) at an auditorily cued 1-Hz pace (alternating with rest epochs with matched auditory stimuli). Only fMRI data from the affected hand were used in this analysis. Follow-up for assessment of ΔFM in the imaged patient sample was at 93 ± 17 days poststroke.

Twelve out of 30 patients of the imaged sample had an acute dynamometry score of 0 kg (Table 1) and thus are likely to have been unable to execute the instructions of the motor task. Formally, however, the experimental condition was not the performance of the task but was rather the instruction to perform the task: behavior is never directly under experimental control and so cannot be properly thought of as an experimental variable. Behavior was not measured, but acute dynamometry score was used as a covariate in the estimation of CV fMRI recovery pattern expression (Z), and so any task-related fMRI activation linearly related to dynamometry could not contribute to prediction of recovery through Z. For example, if task-related activation was simply a reflection of dynamometry score, then Z would be pure noise with respect to ΔFM, even though dynamometry and ΔFM are correlated ($R = 0.56$, 2-tailed $P = 0.001$).

Z was computed in a CV fashion by essentially taking the inner product of a given patient's task-related fMRI activation data with an estimated fMRI recovery pattern whose computation involved neither that patient's fMRI data nor their ΔFM. Z was significantly correlated with ΔFM in the net imaged sample ($R = 0.56$, 1-tailed $P < 0.001$; it is not a typo that this is the same R value as immediately previous). While this result is more robust than a correlation between non-CV recovery pattern expression and ΔFM, it still does not directly indicate the accuracy of prediction. Thus, we also assessed SPE (squared

Table 1
Basic information for imaged sample

Patient ID	Age on date of stroke	Time between stroke and FM_{ii} (days)	Lesion side	Lesion location	FM_{ii}	ΔFM	Acute dynamometry (kg)
1	56	4	R	IC	62	0	0
2	51	1	L	CR/IC	5	5	19
3	59	2	R	Pons	3	3	11
4	70	4	B	Pons	9	6	33
5	48	2	R	Caudate/putamen	3	3	28
6	47	3	L	Pons	42	39	1
7	65	2	R	S1/M1/inferior parietal	41	38	0
8	63	1	L	Thalamus/IC	35	26	0
9	59	2	L	Insula/CR	61	6	0
10	65	2	R	CR	20	15	5
11	77	2	R	Superior parietal/S1/M1	45	34	0
12	67	3	R	IC	52	26	0
13	56	3	R	CR/external capsule	60	0	0
14	59	3	L	Pons	60	40	0
15	42	1	L	Insula/CR	7	7	26
16	69	2	L	Temporal/occipital/pons	29	27	0
17	78	2	B	Occipital/insula/frontal/CR	48	35	0
18	43	2	R	Pons	17	13	30
19	48	1	L	IC	4	3	21
20	72	2	L	Frontal/insula	2	2	16
21	60	2	L	Temporal/parietal/pons	17	15	27
22	45	1	R	CR	4	4	40
23	59	1	L	Pons	58	32	1
24	67	2	R	Midbrain/pons	20	18	11
25	61	3	R	Frontal/temporal/parietal	18	9	23
26	58	2	L	Thalamus/IC	16	12	26
27	72	2	R	Pons	3	1	28
28	72	4	R	IC/CR	2	1	24
29	67	13	B	Medulla/pons	62	28	0
30	55	1	L	Pons	61	−1	0

Note: CR, corona radiata; IC, internal capsule; M1, primary motor cortex; S1, primary sensory cortex; R, right; L, left side; B, bilateral.

prediction error, an explicit measure of prediction accuracy) in the imaged sample for posterior means of ΔFM given FM_{ii} and/or Z. To facilitate appreciation of the magnitude of SPE relative to the range of the FM (0–66), below we will express SPE in the form of "x^2," where x is thus in FM units.

The posterior mean of ΔFM given FM_{ii} required an estimated $f_{\Delta FM|FM_{ii}}$ (which was allowed to be a mixture of K proportional recovery models for each severity category) and $f_{Z|\Delta FM}$. Based on fitting using the nonimaged sample, the minimum AIC $f_{\Delta FM|FM_{ii}}$ for the patients with nonsevere FM_{ii} was $K = 3$ ($\hat{p}_{1,nonsevere}$=0.37, $\hat{\beta}_{1,nonsevere}$=0.55, $\hat{\sigma}^2_{\phi,1,nonsevere}$=27.51, $\hat{p}_{2,nonsevere}$=0.31, $\hat{\beta}_{2,nonsevere}$=0.81, $\hat{\sigma}^2_{\phi,2,nonsevere}$=0.38, $\hat{p}_{3,nonsevere}$=0.32, $\hat{\beta}_{3,nonsevere}$=0.93, $\hat{\sigma}^2_{\phi,3,nonsevere}$=1.22) and the minimum AIC model for the patients with severe FM_{ii} was $K = 1$ ($\hat{p}_{1,severe}$=1, $\hat{\beta}_{1,severe}$=0.44, $\hat{\sigma}^2_{\phi,1,severe}$=376.22). We were surprised that a 1-component model fit best for the severe FM_{ii} group given our previous findings suggesting a mixture of 2 proportional recovery models (one with $\beta \sim 0.7$ and one with $\beta \sim 0$) (Prabhakaran et al. 2008). Visual inspection of the nonimaged sample data confirmed, however, that there was a more even distribution of ΔFM in the patients with severe FM_{ii} from the University of Freiburg (who were not represented in that study) than from Columbia University (who were); we proceeded regardless. The sole parameter of $f_{Z|\Delta FM}$ is σ^2_ϵ. Its maximum likelihood estimator was obtained from Z and ΔFM in the imaged sample: $\hat{\sigma}^2_\epsilon$=644.

A preliminary remark is that it is not a mathematical necessity that SPEs decrease (just as the AIC need not decrease) as more information is added to the prediction algorithm; this can be contrasted with the necessary increase of R^2 for a linear regression model as the rank of the design matrix increases. We now present the (average) SPE in the net imaged sample for the various posterior means of ΔFM: The SPE of the

posterior mean of ΔFM given dichotomous stroke severity was 16^2. The SPE given FM_{ii} was 10^2, while the SPE given FM_{ii} and Z was 8^2. The decrease in SPE in the net imaged sample from using FM_{ii} to using FM_{ii} and Z was not significant ($t(29) = 1.14$, 1-tailed $P = 0.13$).

A more meaningful understanding of the effect of conditioning is provided by looking at SPE separately in the nonsevere ($N = 23$) and severe ($N = 7$) FM_{ii} patient subgroups of the imaged sample (Table 2; Fig. 3). SPE tends to be much greater in all models in the patients with severe FM_{ii}. Relative to conditioning on dichotomous stroke severity alone, conditioning on FM_{ii} improves SPE only in the patients with nonsevere FM_{ii} (nonsevere FM_{ii}: $\Delta SPE = 13^2$, $t(22) = 5.89$, 1-tailed $P < 0.0001$; severe FM_{ii}: $\Delta SPE = -2^2$, $t(6) = -0.66$, 1-tailed $P = 0.73$). Relative to conditioning on FM_{ii}, conditioning on FM_{ii} and Z improves SPE meaningfully, although not significantly, only in the patients with severe FM_{ii} (nonsevere FM_{ii}: $\Delta SPE = 2^2$, $t(22) = 0.98$, 1-tailed $P = 0.17$; severe FM_{ii}: $\Delta SPE = 12^2$, $t(6) = 1.03$, 1-tailed $P = 0.17$).

A way of reexpressing SPE for a given predictive model is in terms of the percentage of the total sum of squares of ΔFM explained by that model, which equals $100 \cdot (1 - [\text{mean SPE for that model}]/[\text{mean } \Delta FM^2])$. Unlike R^2, this value need not be positive. For patients with nonsevere FM_{ii}, this value was 96% and 97% when conditioning on FM_{ii} and [FM_{ii}, Z], respectively. For patients with severe FM_{ii}, this value was 16% and 47% when conditioning on FM_{ii} and [FM_{ii}, Z], respectively.

Discussion

Our opinion is that one focus of stroke research should be the development of novel, early poststroke treatments (Turton and Pomeroy 2002; Biernaskie et al. 2004) for those patients likely

	Severe/nonsevere classification	FM_{ii}	Z and severe/nonsevere classification	FM_{ii} and Z
Net imaged sample (N = 30)	16^2	10^2	13^2	8^2
Patients with severe FM_{ii} (N = 7)	20^2	20^2	16^2	16^2
Patients with nonsevere FM_{ii} (N = 23)	14^2	4^2	12^2	3^2
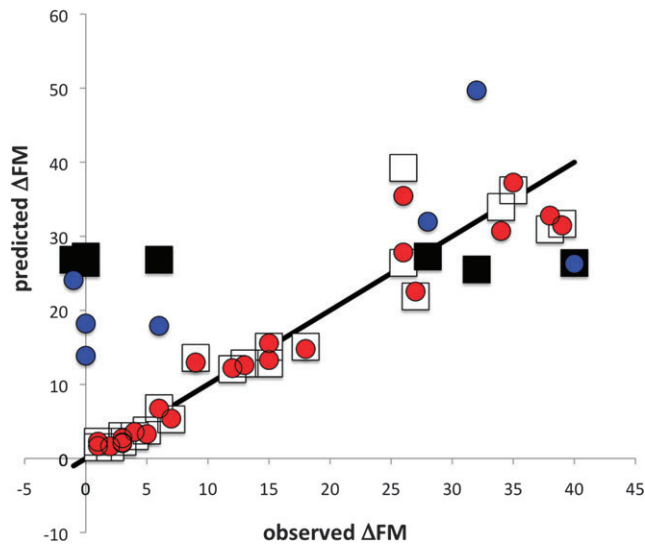
Figure 3. Predicted (posterior mean) versus observed ΔFM when using either 1) FM_{ii} (patients with nonsevere FM_{ii}: white squares; patients with severe FM_{ii}: black squares) or 2) FM_{ii} and Z (patients with nonsevere FM_{ii}: red circles; patients with severe FM_{ii}: blue circles) as the information in determining the posterior density of ΔFM. The SPE for a given datum is the (vertical distance to the identity line)^2. All data plotted correspond to the imaged sample.

to remain severely hemiparetic under current standard care. Such treatments might include noninvasive brain stimulation (Williams et al. 2009) and robotics-based rehabilitation (Huang and Krakauer 2009). However, as such alternative treatments are likely to be more costly in terms of patient effort and healthcare resources than standard care, one ideally wants to exclude patients who would recover substantially under standard care from such alternative treatments. For this purpose, early poststroke prediction should have high sensitivity for recovery. But of course, a high specificity would be desired as well in order to have a low chance of mistakenly excluding patients who will not sufficiently recover. Thus, a generally accurate prediction algorithm would be useful for developing and applying new treatments. An alternative rationale for early prediction of recovery that is more relevant for the current state of treatment is to inform patients and family about expected outcomes and also to direct physical and occupation therapists to focus on compensatory strategies rather than recovery from impairment. A promising prediction finding in this regard is that the presence of finger extension and shoulder abduction 72 h poststroke yielded positive and negative predictive values for some recovery of upper limb dexterity using the action research arm test (ARAT) (for a brief description of the ARAT, which assesses function, see Van der Lee et al. 2001) of 0.98 and 0.75, respectively (Nijland et al.

2010). Here we assessed the contribution of fMRI (through Z) to prediction of recovery from impairment at 2–3 days poststroke. The combination of Z and FM_{ii} led to a non-significant increase from 16% to 47% of the total sum of squares of ΔFM explained in patients with severe FM_{ii}. Assuming the effect size observed here for patients with severe FM_{ii}, 42 patients with severe FM_{ii} would be required for power = 0.80 (58 for power = 0.90; 73 for power = 0.95) to detect such a reduction. However, these sample size calculations are conservative in that as imaged sample size increases, σ_ε^2 would also decrease (and consequently decrease SPE when using Z), while these sample size calculations assume it will remain constant. Likewise, use of a larger sample size to estimate $f_{\Delta FM|FM_{ii}}$ would also tend to lower SPE, while these sample size calculations do not account for such effects. Regardless, we take our negative statistical result with regard to prediction as inconclusive. The significant correlation between Z and ΔFM could be taken as an impetus for sufficiently powered studies. Such studies might also be performed at higher field strength and acquire at larger voxel sizes to increase signal:noise ratio.

Another reason to consider the negative result inconclusive is the nature of the samples used. We estimated $f_{\Delta FM|FM_{ii}}$ from a nonimaged sample combined from 2 institutions. Though the samples were similar in terms of clinical inclusion and exclusion criteria, they were unlikely to be matched genetically, educationally, or socioeconomically (factors which could conceivably affect recovery). We qualitatively observed (after the decision had been made to use the combined sample) that the patients with severe initial impairment from the University of Freiburg had a less stark dichotomy of ΔFM than those from Columbia University. This made $f_{\Delta FM|FM_{ii}}$ for this severity a single broad Gaussian rather than 2 well-separated, narrower ones (the latter having been our hypothesis based on the results of a previous study; Prabhakaran et al. 2008). As the imaged sample (exclusively from Columbia University) manifested dichotomous ΔFM, a speculative explanation for the unimpressive performance of the posterior mean based on Z and FM_{ii} is that $f_{\Delta FM|FM_{ii}}$ was too far from ground truth for the population from which the imaged sample was obtained. In hindsight, we recognize that we were overeager to expand our sample size for estimating $f_{\Delta FM|FM_{ii}}$ and consequently did not consider the possibility of important systematic differences in populations sampled by the 2 institutions; we hope to be more careful in our future work.

Current Level of Impairment and Subsequent Recovery Could Have Different Neural Correlates

While the purpose here was simply to predict ΔFM from FM_{ii} and fMRI, rather than characterize the brain mechanisms of recovery, it is important to make clear how the experimental design we employed relates to those of previous relevant brain

imaging studies of stroke. Previous such studies can be dichotomized into longitudinal (Marshall et al. 2000; Calautti et al. 2001; Feydy et al. 2002; Ward et al. 2003b,) and cross-sectional (Ward et al. 2003a; Loubinoux et al. 2003; Loubinoux et al. 2003; Loubinoux et al. 2007; Jang et al. 2004) designs; the design of the current study was cross-sectional. Ward et al. (2003a, 2004) have argued that the existence of positive cross-sectional correlations between activation and impairment at the time of scanning in both the early and chronic periods indicates that patients with more severe impairment need to utilize "secondary motor areas" to a greater extent than less impaired patients in order to generate motor output (which is nevertheless suboptimal). This hypothesis is supported by a study that showed that the slowing of reaction time induced by TMS applied to contralesional dorsal premotor cortex correlates with the degree of impairment (Johansen-Berg et al. 2002). However, it is not clear how the cross-sectional correlations observed by Ward et al. (2003a, 2004) relate, either in a causal sense or simply in a correlative sense, to subsequent recovery (i.e., recovery that manifests behaviorally after the time of imaging). Spatial signatures for these 2 types of theoretical activation signals (considered at any fixed time poststroke), those related to current level of impairment (ostensibly reflected in the results of Ward et al. 2003a, 2004) and those related to subsequent recovery (ostensibly contributing to the current prediction results and similarly to the correlations from our previous imaging study), need not be the same. This is because the failure of proportional recovery in patients with severe initial impairment uncouples current level of impairment and subsequent recovery; conversely, in the absence of patients with severe initial impairment who recover very little, it would not be possible to disambiguate the neural correlates of current level of impairment and subsequent recovery. It is therefore of note that the studies of Ward et al. (2003b, 2004) seem not to have included any such patients, which implies that their reported correlations from the early period could be an admixture of those related to current level of impairment and those related to subsequent recovery. This could possibly be an explanation of their finding that certain correlations detected early were not detected later poststroke (the ones only detected early possibly being related to subsequent recovery) (Ward et al. 2003b, 2004).

Three cross-sectional studies have concerned correlation and/or prediction of future impairment level (as opposed to change in impairment) (Loubinoux et al. 2003; Jang et al. 2004; Loubinoux et al. 2007). Two of these studies from the same group were correlative (i.e., not predictive) and yielded inconsistent results with respect to one another (Loubinoux et al. 2003, 2007). The third study (Jang et al. 2004) attempted prospective prediction based on an arbitrary decision criterion applied to voxel-wise activation data in a primary sensorimotor cortical region of interest; the resulting prediction accuracy was poor. This type of approach was not devised to be optimal given the region of interest chosen, nor did it reasonably weight all potential spatial sources of relevant functional imaging signal (Kjems et al. 2002; Strother et al. 2002; O'Toole et al. 2007). In contrast, the approach taken in the current study was to use the posterior mean (which is an optimal estimator given a conditional density) and also to use all the linear information about ΔFM present in multivariate fMRI data (not just that from a single region of interest). However, neither study yielded impressive prediction.

MEPs and Recovery in Patients with Severe Initial Impairment

The volume of anatomical damage per se seems insufficient to explain the majority of the variation in recovery in patients with severe initial impairment (Binkofski et al. 2001; Shelton and Reding 2001; Konishi et al. 2005; Cho et al. 2007). MEPs ostensibly measure the functional nature of the damage sustained to the CST (van Kuijk et al. 2005), as opposed to the total lesion volume. Pooling data from 3 studies that examined MEPs in hemiplegic patients within the first 10 days poststroke (Hendricks et al. 1997; Hendricks, Pasman, Merx et al. 2003; van Kuijk et al. 2009), positive predictive value for recovery was 0.94 and negative predictive value was 0.83. Thus, while it does seem that the degree of functional CST damage is a critical determinant of recovery, recovery is possible in 15–20% of cases where functional CST damage seems complete; this recovery could depend on brain plasticity. Moreover, patients MEP$^+$ at 1 week tend not to reach their recovery endpoints until 1–3 months later (Swayne et al. 2008), which suggests that even in the presence of residual monosynaptic connections from primary motor cortex (M1) to spinal motoneurons, some sort of plasticity mechanism which lasts for several weeks poststroke is required to allow their effective use. fMRI could be sensitive to these putative brain plasticity mechanisms. It would therefore be of interest to specify a conditional density (or better, several densities representing different hypotheses) for recovery given initial impairment, fMRI, and MEPs.

Caveats

No behavioral measure was acquired during imaging. Therefore, the correlation between Z and ΔFM could be due to individual differences in behavior during imaging. Such behavior can include, but is not limited to, mirror movements, which have been estimated to occur 70% of the time during repetitive squeezes using the affected hand and correlate with severity of impairment of the affected hand (Nelles et al. 1998). Activation of contralesional motor cortex correlates with individual differences in mirror movements in hemiparetic stroke patients (Wittenberg et al. 2000; Kim et al. 2003). However, even if individual differences in the degree of mirror movements during imaging were causal to the observed correlation between Z and ΔFM, it would not invalidate prediction of ΔFM based on fMRI using the current design. This is because the metric for prediction is the magnitude of prediction error; nothing more, nothing less (Akaike 1974). While it is true that the inclusion of more controls at the experimental design (e.g., clamp performance in the scanner) and/or analysis stages (e.g., include a covariate for performance in the scanner) could potentially improve prediction by reducing the variance of $\hat{d}$, it is logically incorrect to say that a degree of prediction obtained in the absence of such controls "does not count" or is "artificially too high." This is an important difference between predictive and causal modeling, and this was a prediction study.

If the correlation between Z and ΔFM were caused only by correlations between 1) FM_{ii} and the degree of mirror movements (Nelles et al. 1998), 2) mirror movements and contralesional M1 activation (Wittenberg et al. 2000; Kim et al. 2003), and 3) FM_{ii} and ΔFM (Prabhakaran et al. 2008), then (assuming the measurement noise of FM_{ii} is negligible) Z would not improve the prediction of ΔFM over that provided by FM_{ii}

alone. The empirical result was that while recovery pattern expression was significantly correlated with recovery, its contribution to the prediction of recovery over that provided by initial impairment alone was not statistically significant. Thus, the result is on its face consistent with mirror movements having contributed to recovery pattern expression. In future studies, mirror movements could be measured quantitatively and included as a covariate in recovery pattern estimation (see fMRI-based Measurement of ΔFM); this would eliminate any (linear) contribution of mirror movements to Z. Doing so could 1) reduce unexplained variability in the fMRI data (good for prediction, as it *ceteris paribus* decreases the variance of the recovery pattern $\hat{d}$ and hence decreases σ_ε^2), 2) reduce the amount of variation in the fMRI data uniquely attributable to ΔFM (bad for prediction, as it *ceteris paribus* increases variance of $\hat{d}$ and hence increases σ_ε^2), and 3) increase the complexity of the model by adding more parameters (which is bad for prediction, as it *ceteris paribus* increases σ_ε^2). The net effect on the accuracy of prediction of ΔFM would depend on the balance of these effects.

It is possible that some patients did not perform the instructed movement with the affected hand at all during imaging. Indeed, 12/30 patients in the imaged sample had a grip force dynamometry score (of the affected hand) of 0 on the day of scanning; recovery was heterogeneous in this subgroup (Table 1). In the prediction algorithm, both the estimation of **d** and the computation of Z use dynamometry score as a covariate such that any systematic component of the fMRI signal that is linearly dependent on grip force dynamometry has no effect on the prediction of ΔFM. Thus, if the intensity of movement (say the across-click average of the maximal torque produced by the effector per metronome click) during scanning is linearly related to dynamometry score, then whether subjects actually moved or not is irrelevant to prediction of ΔFM in this model. If instead, the intensity of movement during scanning is only weakly correlated with dynamometry score (or not correlated at all) and if the pattern of brain activation associated with intensity of movement is strongly "spatially" correlated with the recovery pattern, then this would reduce the signal:noise of Z (as the random, unmodeled variations across subjects in intensity of movement during scanning would lead to activations similar to those associated with recovery and hence add noise to Z) and hence worsen prediction. If this were the case, then measurement of the intensity of movement within the scanner would allow for improved prediction. A third possibility is that brain activation correlates of recovery are modulated by the intensity of movement; if so, failure (as in the current study) to account for this modulation (i.e., interaction) would lead to a worsened prediction. It is again worth noting that there is no way that failure to account for movement (or anything else) could lead to an artificially good prediction of ΔFM.

The patients in the imaged sample had predominantly subcortical strokes. It is an empirical question as to whether fMRI would contain predictive information about recovery for large cortical strokes. If the cortical regions damaged were critical for recovery, then perhaps this would be reflected in lower Z (in which case, the predictive information of Z would be preserved).

Conclusion

In conclusion, we found that prediction of recovery in patients with nonsevere initial impairment was accurate based on initial impairment alone. In contrast, prediction of recovery in patients with severe initial impairment was poor based on initial impairment alone and was not statistically significantly improved by the inclusion of fMRI acquired at 2 days poststroke. However, the significant correlation between fMRI and recovery might provide a motivation for further assessment of prediction using adequately powered studies and more carefully considered samples.

Supplementary Material

Supplementary material can be found at: http://www.cercor .oxfordjournals.org/

References

Akaike H. 1974. New look at statistical-model identification. IEEE Trans Automat Contr. Ac19:716–723.

Biernaskie J, Chernenko G, Corbett D. 2004. Efficacy of rehabilitative experience declines with time after focal ischemic brain injury. J Neurosci. 24:1245–1254.

Binkofski F, Seitz RJ, Hacklander T, Pawelec D, Mau J, Freund HJ. 2001. Recovery of motor functions following hemiparetic stroke: a clinical and magnetic resonance-morphometric study. Cerebrovasc Dis. 11:273–281.

Burnham KP, Anderson DR. 2002. Model selection and multimodel inference: a practical information-theoretic approach. New York (NY): Springer-Verlag.

Calautti C, Leroy F, Guincestre JY, Baron JC. 2001. Dynamics of motor network overactivation after striatocapsular stroke: a longitudinal PET study using a fixed-performance paradigm. Stroke. 32:2534–2542.

Cho SH, Kim SH, Choi BY, Kang JH, Lee CH, Byun WM, Jang SH. 2007. Motor outcome according to diffusion tensor tractography findings in the early stage of intracerebral hemorrhage. Neurosci Lett. 421:142–146.

Duncan PW, Goldstein LB, Horner RD, Landsman PB, Samsa GP, Matchar DB. 1994. Similar motor recovery of upper and lower extremities after stroke. Stroke. 25:1181–1188.

Duncan PW, Goldstein LB, Matchar D, Divine GW, Feussner J. 1992. Measurement of motor recovery after stroke. Outcome assessment and sample size requirements. Stroke. 23:1084–1089.

Duncan PW, Propst M, Nelson SG. 1983. Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. Phys Ther. 63:1606–1610.

Efron B, Tibshirani R. 1981. Statistical data analysis in the computer age. Science. 253:390–395.

Feydy A, Carlier R, Roby-Brami A, Bussel B, Cazalis F, Pierot L, Burnod Y, Maier MA. 2002. Longitudinal study of motor recovery after stroke: recruitment and focusing of brain activation. Stroke. 33: 1610–1617.

Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. 1975. The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. Scand J Rehabil Med. 7:13–31.

Gladstone DJ, Danells CJ, Black SE. 2002. The Fugl-Meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil Neural Repair. 16:232–240.

Gowland C, Stratford P, Ward M, Moreland J, Torresin W, Van Hullenaar S, Sanford J, Barreca S, Vanspall B, Plews N. 1993. Measuring physical impairment and disability with the Chedoke-McMaster Stroke Assessment. Stroke. 24:58-63.

Hendricks HT, Hageman G, van Limbeek J. 1997. Prediction of recovery from upper extremity paralysis after stroke by measuring evoked potentials. Scand J Rehabil Med. 29:155-159.

Hendricks HT, Pasman JW, Merx JL, van Limbeek J, Zwarts MJ. 2003. Analysis of recovery processes after stroke by means of transcranial magnetic stimulation. J Clin Neurophysiol. 20:188-195.

Hendricks HT, Pasman JW, van Limbeek J, Zwarts MJ. 2003. Motor evoked potentials in predicting recovery from upper extremity paralysis after acute stroke. Cerebrovasc Dis. 16:265-271.

Hendricks HT, van Limbeek J, Geurts AC, Zwarts MJ. 2002. Motor recovery after stroke: a systematic review of the literature. Arch Phys Med Rehabil. 83:1629-1637.

Huang VS, Krakauer JW. 2009. Robotic neurorehabilitation: a computational motor learning perspective. J Neuroeng Rehabil. 6:5.

Jang SH, Kim YH, Chang Y, Han BS, Byun WM, Chang CH. 2004. The predictive value of cortical activation by passive movement for motor recovery in stroke patients. Restor Neurol Neurosci. 22:59-63.

Johansen-Berg H, Rushworth MF, Bogdanovic MD, Kischka U, Wimalaratna S, Matthews PM. 2002. The role of ipsilateral premotor cortex in hand movement after stroke. Proc Natl Acad Sci U S A. 99:14518-14523.

Jorgensen HS, Nakayama H, Raaschou HO, Olsen TS. 1999. Stroke. Neurologic and functional recovery: the Copenhagen stroke study. Phys Med Rehabil Clin N Am. 10:887-906.

Jorgensen HS, Nakayama H, Raaschou HO, Vive-Larsen J, Stoier M, Olsen TS. 1995. Outcome and time course of recovery in stroke. Part II: time course of recovery. The Copenhagen Stroke Study. Arch Phys Med Rehabil. 76:406-412.

Kim YH, Jang SH, Chang Y, Byun WM, Son S, Ahn SH. 2003. Bilateral primary sensori-motor cortex activation of post-stroke mirror movements: an fMRI study. Neuroreport. 14:1329-1332.

Kjems U, Hansen LK, Anderson J, Frutiger S, Muley S, Sidtis J, Rottenberg D, Strother SC. 2002. The quantitative evaluation of functional neuroimaging experiments: mutual information learning curves. Neuroimage. 15:772-786.

Konishi J, Yamada K, Kizu O, Ito H, Sugimura K, Yoshikawa K, Nakagawa M, Nishimura T. 2005. MR tractography for the evaluation of functional recovery from lenticulostriate infarcts. Neurology. 64:108-113.

Krakauer JW. 2005. Arm function after stroke: from physiology to recovery. Semin Neurol. 25:384-395.

Kwakkel G, Kollen B, Lindeman E. 2004. Understanding the pattern of functional recovery after stroke: facts and theories. Restor Neurol Neurosci. 22:281-299.

Kwakkel G, Kollen B, Twisk J. 2006. Impact of time on improvement of outcome after stroke. Stroke. 37:2348-2353.

Kwakkel G, Kollen BJ, van der Grond J, Prevo AJ. 2003. Probability of regaining dexterity in the flaccid upper limb: impact of severity of paresis and time since onset in acute stroke. Stroke. 34:2181-2186.

Loubinoux I, Carel C, Pariente J, Dechaumont S, Albucher JF, Marque P, Manelfe C, Chollet F. 2003. Correlation between cerebral re-organization and motor recovery after subcortical infarcts. Neuroimage. 20:2166-2180.

Loubinoux I, Dechaumont-Palacin S, Castel-Lacanal E, De Boissezon X, Marque P, Pariente J, Albucher JF, Berry I, Chollet F. 2007. Prognostic value of FMRI in recovery of hand function in subcortical stroke patients. Cereb Cortex. 17:2980-2987.

Lyle RC. 1981. A performance test for assessment of upper limb function in physical rehabilitation treatment and research. Int J Rehabil Res. 4:483-492.

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

Marshall RS, Zarahn E, Alon L, Minzer B, Lazar RM, Krakauer JW. 2009. Early imaging correlates of subsequent motor recovery after stroke. Ann Neurol. 65:596-602.

Nakayama H, Jorgensen HS, Raaschou HO, Olsen TS. 1994. Compensation in recovery of upper extremity function after stroke: the Copenhagen Stroke Study. Arch Phys Med Rehabil. 75:852-857.

Nelles G, Cramer SC, Schaechter JD, Kaplan JD, Finklestein SP. 1998. Quantitative assessment of mirror movements after stroke. Stroke. 29:1182-1187.

Nijland RH, van Wegen EE, Harmeling-van der Wel BC, Kwakkel G. 2010. Presence of finger extension and shoulder abduction within 72 hours after stroke predicts functional recovery: early prediction of functional outcome after stroke: the EPOS cohort study. Stroke. 41:745-750.

O'Toole AJ, Jiang F, Abdi H, Penard N, Dunlop JP, Parent MA. 2007. Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. J Cogn Neurosci. 19:1735-1752.

Platz T, Pinkowski C, van Wijck F, Kim IH, di Bella P, Johnson G. 2005. Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: a multicentre study. Clin Rehabil. 19:404-411.

Prabhakaran S, Zarahn E, Riley C, Speizer A, Chong JY, Lazar RM, Marshall RS, Krakauer JW. 2008. Inter-individual variability in the capacity for motor recovery after ischemic stroke. Neurorehabil Neural Repair. 22:64-71.

Shao J. 2003. Mathematical statistics. New York (NY): Springer.

Shelton FN, Reding MJ. 2001. Effect of lesion location on upper limb motor recovery after stroke. Stroke. 32:107-112.

Stone M. 1974. Cross-validatory choice and assessment of statistical predictions. J R Stat Soc Series B Methodol. 36:111-147.

Stone M. 1977. An asymptotic equivalence of choice of model by cross-validation and Akaikes criterion. J R Stat Soc Series B Stat Methodol. 39:44-47.

Strother SC, Anderson J, Hansen LK, Kjems U, Kustra R, Sidtis J, Frutiger S, Muley S, LaConte S, Rottenberg D. 2002. The quantitative evaluation of functional neuroimaging experiments: the NPAIRS data analysis framework. Neuroimage. 15:747-771.

Swayne OB, Rothwell JC, Ward NS, Greenwood RJ. 2008. Stages of motor output reorganization after hemispheric stroke suggested by longitudinal studies of cortical physiology. Cereb Cortex. 18:1909-1922.

Turton A, Pomeroy V. 2002. When should upper limb function be trained after stroke? Evidence for and against early intervention. Neurorehabilitation. 17:215-224.

Van der Lee JH, De Groot V, Beckerman H, Wagenaar RC, Lankhorst GJ, Bouter LM. 2001. The intra- and interrater reliability of the action research arm test: a practical test of upper extremity function in patients with stroke. Arch Phys Med Rehabil. 82:14-19.

van der Lee JH, Wagenaar RC, Lankhorst GJ, Vogelaar TW, Deville WL, Bouter LM. 1999. Forced use of the upper extremity in chronic stroke patients: results from a single-blind randomized clinical trial. Stroke. 30:2369-2375.

van Kuijk AA, Pasman JW, Geurts AC, Hendricks HT. 2005. How salient is the silent period? The role of the silent period in the prognosis of upper extremity motor recovery after severe stroke. J Clin Neurophysiol. 22:10-24.

van Kuijk AA, Pasman JW, Hendricks HT, Zwarts MJ, Geurts AC. 2009. Predicting hand motor recovery in severe stroke: the role of motor evoked potentials in relation to early clinical assessment. Neurorehabil Neural Repair. 23:45-51.

Verheyden G, Nieuwboer A, De Wit L, Thijs V, Dobbelaere J, Devos H, Severijns D, Vanbeveren S, De Weerdt W. 2008. Time course of trunk, arm, leg, and functional recovery after ischemic stroke. Neurorehabil Neural Repair. 22:173-179.

Ward NS, Brown MM, Thompson AJ, Frackowiak RS. 2003a. Neural correlates of outcome after stroke: a cross-sectional fMRI study. Brain. 126(Pt 6):1430-1448.

Ward NS, Brown MM, Thompson AJ, Frackowiak RS. 2003b. Neural correlates of motor recovery after stroke: a longitudinal fMRI study. Brain. 126:2476-2496.

Ward NS, Brown MM, Thompson AJ, Frackowiak RS. 2004. The influence of time after stroke on brain activations during a motor task. Ann Neurol. 55:829-834.

Williams JA, Imamura M, Fregni F. 2009. Updates on the use of non-invasive brain stimulation in physical and rehabilitation medicine. J Rehabil Med. 41:305-311.

Wittenberg GF, Bastian AJ, Dromerick AW, Thach WT, Powers WJ. 2000. Mirror movements complicate interpretation of cerebral activation changes during recovery from subcortical infarction. Neurorehabil Neural Repair. 14:213-221.

Wolfe CD. 2000. The impact of stroke. Br Med Bull. 56:275-286.

Woodbury ML, Velozo CA, Richards LG, Duncan PW, Studenski S, Lai SM. 2008. Longitudinal stability of the Fugl-Meyer Assessment of the upper extremity. Arch Phys Med Rehabil. 89:1563–1569.