

Individual Prediction of Chronic Motor Outcome in the Acute Post-Stroke Stage: Behavioral Parameters Versus Functional Imaging

Anne K. Rehme,^{1,2*} Lukas J. Volz,^{1,3} Delia-Lisa Feis,⁴ Simon B. Eickhoff,^{2,5}
Gereon R. Fink,^{1,2} and Christian Grefkes^{1,2}

¹Department of Neurology, University of Cologne, Cologne, Germany

²Institute of Neuroscience and Medicine (INM-1, INM-3), Research Centre Jülich, Jülich, Germany

³Department of Psychological and Brain Sciences, University of California, Santa Barbara, USA

⁴Max Planck Institute for Metabolism Research, Cologne, Germany

⁵Institute of Clinical Neuroscience and Medical Psychology, Heinrich Heine University, Düsseldorf, Germany

Abstract: Several neurobiological factors have been found to correlate with functional recovery after brain lesions. However, predicting the individual potential of recovery remains difficult. Here we used multivariate support vector machine (SVM) classification to explore the prognostic value of functional magnetic resonance imaging (fMRI) to predict individual motor outcome at 4–6 months post-stroke. To this end, 21 first-ever stroke patients with hand motor deficits participated in an fMRI hand motor task in the first few days post-stroke. Motor impairment was quantified assessing grip force and the Action Research Arm Test. Linear SVM classifiers were trained to predict good versus poor motor outcome of unseen new patients. We found that fMRI activity acquired in the first week post-stroke correctly predicted the outcome for 86% of all patients. In contrast, the concurrent assessment of motor function provided 76% accuracy with low sensitivity (<60%). Furthermore, the outcome of patients with initially moderate impairment and high outcome variability could not be predicted based on motor tests. In contrast, fMRI provided 87.5% prediction accuracy in these patients. Classifications were driven by activity in ipsilesional motor areas and contralesional cerebellum. The accuracy of subacute fMRI data (two weeks post-stroke), age, time post-stroke, lesion volume, and location were at 50%-chance-level. In conclusion, multivariate decoding of fMRI data with SVM early after stroke enables a robust prediction of motor recovery. The potential for recovery is influenced by the initial dysfunction of the active motor system, particularly in those patients whose outcome cannot be predicted by behavioral tests. *Hum Brain Mapp* 36:4553–4565, 2015. © 2015 Wiley Periodicals, Inc.

Additional Supporting Information may be found in the online version of this article.

Contract grant sponsor: German Research Foundation (DFG); Contract grant numbers: EI 816/4-1; EI 816/6-1; LA 3071/3-1; GR 3285/2-1; GR 3285/4-1; Contract grant sponsor: National Institute of Mental Health; Contract grant number: R01-MH074457; Contract grant sponsor: European EFT program (Human Brain Project), University of Cologne Emerging Groups Initiative (CONNECT group) implemented into the Institutional Strategy of the University of Cologne and the German Excellence Initiative; Marga and Walter Boll Stiftung

*Correspondence to: Dr. Anne K. Rehme, PhD, Department of Neurology, University of Cologne, Kerpener Str. 62, 50924 Germany. E-mail: anne.rehme@uk-koeln.de

Received for publication 17 February 2015; Revised 15 July 2015; Accepted 3 August 2015.

DOI: 10.1002/hbm.22936

Published online 19 August 2015 in Wiley Online Library (wileyonlinelibrary.com).

Key words: functional magnetic resonance imaging (fMRI); motor impairment; multivariate pattern classification; motor recovery; support vector machine

INTRODUCTION

Stroke lesions cause permanent motor disability in some patients whereas others recover substantially within the first days and weeks post-stroke [Go et al., 2013; Miller et al., 2010]. Prognostic regression models demonstrated that the severity of the acute motor impairment predicts chronic motor outcome. Specifically, those patients presenting initially with mild deficits usually reveal good motor outcome while patients with severe deficits often remain chronically disabled [Kwakkel and Kollen, 2007; Nijland et al., 2010; Veerbeek et al., 2011]. In contrast, prognosis becomes rather complicated for the broad spectrum of patients with initially moderate deficits that typically show high variability in neurological outcome [Prabhakaran et al., 2008; Stinear, 2010; Stinear et al., 2012; Zarahn et al., 2011]. The prospective outcome of these patients is, accordingly, more difficult to predict on the basis of initial behavioral measures.

The variability in chronic motor impairment after stroke has been linked to structural and functional changes in brain networks [Cramer, 2008]. For example, the integrity of the corticospinal tract (CST) has been shown to predict motor outcome after stroke [Hendricks et al., 2002; Stinear et al., 2007, 2012; Swayne et al., 2008]. Likewise, brain activity measured with functional magnetic resonance imaging (fMRI) has been suggested to serve as a prognostic marker of chronic disability in addition to behavioral and anatomical measures [Cramer et al., 2007; Rehme and Grefkes, 2013; Stinear and Ward, 2013]. A number of fMRI studies demonstrated that motor recovery after stroke is associated with a reinstatement of previously reduced neural activity in ipsilesional primary motor areas during movements of the paretic upper limb [Cramer et al., 2007; Loubinoux et al., 2007; Rehme et al., 2011b; Tombari et al., 2004; Ward et al., 2003]. Furthermore, regression analyses showed that early motor-related fMRI activity patterns correlated with subsequent motor recovery independent of initial motor impairment [Cramer et al., 2007; Marshall et al., 2009; Zarahn et al., 2011].

To date, clinically valid prognostic models for stroke rehabilitation are scarce because most studies were based on differences or correlations at the group level but provide no information on individual outcome [Orrù et al., 2012]. However, the variability in clinical recovery calls for individualized models in order to enable clinical decision-making based on individual predictors [Stinear, 2010]. The advent of statistical learning models in neuroimaging may help to decode phenotypical information from brain activation patterns at the level of individual patients [Nasela-

ris et al., 2011; Orrù et al., 2012]. For example, support vector machines (SVM) can be used to decode individual motor impairment in acute stroke based on resting-state fMRI [Rehme et al., in press]. Likewise, the language outcome of stroke patients with aphasia can be predicted based on SVM analyses of early fMRI activation during a language task [Saur et al., 2010]. These results lead to the question whether a similar approach allows prediction of individual motor outcome after stroke.

We here tested whether fMRI in the first few days after stroke predicts chronic motor outcome at the level of individual patients. In particular, we aimed to examine whether fMRI can improve the accuracy of the prognosis of patients over and above the accuracy provided by the assessment of motor function. To this end, we assessed motor behavior and acquired task-related fMRI data during voluntary movements of the affected hand in the first and second week post-stroke in a sample of $n = 21$ stroke patients. We trained linear SVM classifiers to predict good versus poor motor outcome at the chronic stage. We hypothesized that fMRI data provide a more accurate individual prognosis than motor impairment alone. Specifically, we expected fMRI data to show a greater predictive value than motor behavior in patients with moderate initial motor impairment as these patients show a great variability in their motor outcome [Prabhakaran et al., 2008; Stinear et al., 2012; Zarahn et al., 2011]. We further assumed that activity in ipsilesional motor areas represents a prognostic marker of chronic motor outcome [Cramer et al., 2007; Loubinoux et al., 2007; Rehme et al., 2011b; Tombari et al., 2004; Ward et al., 2003].

MATERIALS AND METHODS

Sample

We recruited 21 acute stroke patients (mean age: 67 years \pm 11 years, range: 42–89 years, 16 males; Table 1) from the Stroke Unit of the Department of Neurology, University of Cologne, Germany. Inclusion criteria were: (1) first-ever ischemic stroke in the left or right middle cerebral artery territory, (2) unilateral hand motor deficit, (3) symptom onset <10 days, and (4) no clinical signs of mirror movements. Exclusion criteria were: (1) hemorrhagic stroke, (2) bilateral stroke lesions, (3) contraindications to MRI, (4) cognitive impairment or dementia, (5) impaired consciousness, and (6) other orthopedic, neurological, or psychiatric disease.

Behavioral and neuroimaging data were assessed in all $n = 21$ patients at two time-points during the post-stroke

TABLE I. Clinical data of stroke patients

Patient no.	Age	Gender	Time since stroke (days)	Lesion volume (mm ³)	Percent CST damage (from DWI)	Exp.	Acute stage		Subacute stage		Chronic stage		Composite motor score (chronic stage)	Outcome Group (1 = poor outcome 2 = good outcome)
							ARAT score	Grip force index (%)	ARAT score	Grip force index (%)	ARAT score	Grip force index (%)		
Patient 01	84	f	88	12,214	4.2	1	0	0	11	22	53	68	0.68	1
Patient 02	71	m	143	44,452	10.4	1	0	0	0	0	0	0	0.00	1
Patient 03	68	m	272	2,039	4	1	43	86	44	98	57	55	0.68	1
Patient 04	59	f	118	3,945	6.1	1	0	0	0	10	0	10	0.03	1
Patient 05	64	m	101	6,058	2.2	1	28	38	49	49	56	59	0.68	1
Patient 06	65	m	178	6,672	15.1	2	0	0	0	0	0	17	0.05	1
Patient 07	59	m	147	21,664	9.8	2	0	0	0	0	35	17	0.36	1
Patient 08	42	m	132	114,270	23.1	2	38	44	46	65	48	79	0.67	1
Patient 09	89	m	114	3,797	18.1	2	43	44	57	47	55	56	0.66	1
Patient 10	83	f	105	3,459	7.2	2	13	37	32	113	57	61	0.69	1
Patient 11	64	m	131	577	2.4	1	50	63	57	72	57	66	0.71	2
Patient 12	59	m	125	1,640	2.9	1	54	74	56	86	57	75	0.74	2
Patient 13	53	m	177	4,084	4.6	1	54	112	57	108	57	99	0.82	2
Patient 14	77	f	120	13,932	0.1	1	53	86	53	90	53	90	0.75	2
Patient 15	61	m	134	2,926	0.2	1	55	87	57	94	57	106	0.84	2
Patient 16	69	f	111	1,846	42.4	1	38	94	41	100	47	113	0.77	2
Patient 17	59	m	118	3,395	8.5	1	12	7	36	97	56	73	0.73	2
Patient 18	71	m	153	52,805	10.4	2	49	25	56	48	57	156	1.00	2
Patient 19	72	m	99	11,434	1.2	2	32	33	43	36	57	64	0.71	2
Patient 20	72	m	134	21,347	0.5	2	56	90	57	93	57	111	0.86	2
Patient 21	75	f	141	42,430	3.2	2	37	63	50	48	57	112	0.86	2
Mean	67		135	17,856	7.9		31	47	38	61	46	71	median = 0.71	
SD	11		39	26,922	10.1		22	37	22	38	20	39		

ARAT: Action Research Arm Test (affected hand; maximum score = 57); CST: corticospinal tract; DWI: diffusion weighted imaging; f: female; m: male. Time after stroke refers to the days from stroke onset to the chronic stage.

recovery phase. The first session was at the end of the first week (6 ± 3 days post-stroke), here referred to as the *acute stage* after stroke, when numerous processes around the lesion and in remote areas occur as a result of the tissue damage [Carmichael, 2003; Cramer, 2008]. Patients participated in a second session after two weeks (12 ± 4 days post-stroke). This time period is referred to as the *subacute stage* [Cramer, 2008; Duncan et al., 1992; Kwakkel and Kollen, 2007]. Finally, all patients participated in a behavioral follow-up assessment after 4–6 months (135 ± 39 days post-stroke). This late period marks the beginning of the *chronic stage* when neural processes and behavioral improvements begin to stagnate and deficits become permanent [Cramer, 2008; Duncan et al., 1992; Kwakkel and Kollen, 2007]. All patients received standard physiotherapy.

The fMRI data were combined from two experiments with similar setting and course described below which were both approved by the local ethics committee (File No. 08-082 and 09-108). Informed consent was obtained from each subject. Data of the first experiment ($n = 12$) were from a previous study on cortical reorganization after stroke [Rehme et al., 2011b]. In this study, conventional univariate General Linear Model (GLM) analyses revealed

group differences between severely and mildly affected patients or healthy controls in brain activity of ipsilesional and contralesional motor areas during the time course of recovery. In the present study, we used an entirely different statistical approach consisting of multivariate SVM pattern classification to test whether regional brain activity allows predicting outcome of individual patients [Rehme et al., in press]. Data from the second experiment ($n = 9$) stem from an ongoing study on longitudinal recovery after motor stroke (Table I). Using χ^2 - and t -tests in SPSS, we found no difference in any clinical parameter including age, gender, motor impairment, time after stroke onset, and CST damage between the two samples (all P -values ≥ 0.2). Thus, the combination of data from different experiments was unlikely to impact on the SVM results.

Magnetic Resonance Imaging

Data of both experiments were acquired on the same 3 T scanner (Siemens Medical Solutions, Erlangen, Germany). In both experiments, patients performed blocks of unimanual movements consisting of visually cued rhythmic fist

closures at a frequency of 1 Hz. Visual instructions were presented on a shielded thin-film transistor screen at the rear end of the scanner, which was visible via a mirror at the MRI head coil. Written instructions were displayed for 2 s on a screen indicating whether the left or right hand had to be moved in the upcoming block of trials. Each block lasted about 15 s during which subjects moved one hand according to the rhythm (1 Hz) of a blinking circle until a black screen indicated to rest for 15 s (plus a temporal jitter of 1–2.5 s). Prior to scanning, subjects were familiarized with the task until they reached stable performance in three successive trials. Patients who could not achieve the requested frequency were instructed to perform correct fist closures as close as possible to the visual cue.

Both experiments were measured using a gradient echo-planar imaging MRI sequence with slices covering the whole brain from the vertex to lower parts of the cerebellum. In the first experiment, the imaging parameters were as follows: time of repetition (TR) = 1,630 ms, echo time (TE) = 30 ms, field of view (FOV) = 200 mm, 26 axial slices, voxel size = $3.0 \times 3.0 \times 4.0 \text{ mm}^3$, flip angle = 72° , volumes = 176 (4 dummy scans). In the second experiment, the following parameters were used: TR = 2,200 ms, TE = 30 ms, FOV = 200 mm, 33 slices, voxel size: $3.1 \times 3.1 \times 3.1 \text{ mm}^3$, flip angle = 90° , 283 volumes (4 dummy scans). Patients from both samples were equally distributed in the two outcome groups of the SVM analysis ($\chi^2(1) = 0.398$; $P > 0.5$). Furthermore, we computed χ^2 -tests to test whether SVM classifications were driven by any difference in MRI sequence parameters. At each session, we acquired diffusion-weighted images (DWI; TR = 5,100 ms, TE = 104 ms, FOV = 230 mm, 30 axial slices, voxel size = $1.8 \times 1.8 \times 3.0 \text{ mm}^3$) to assess location and extent of the acute ischemic lesion.

fMRI Analysis

The fMRI time-series were analyzed using statistical parametric mapping (SPM8; <http://www.fil.ion.ucl.ac.uk/spm/software/spm8/>). Prior to data analysis, fMRI data of patients with right-hemispheric lesions were flipped along the midsagittal plane. Thus, after flipping, the left hemisphere corresponded to the affected hemisphere and the right hand represented the affected hand in all patients. All patients ($n = 21$) participated at two fMRI sessions at the acute and subacute stage. These two data sets were processed in independent analyses in order to test whether prediction accuracy differs between both stages. The first four images of each time-series were discarded. The remaining images were spatially realigned to the mean image of the time-series and co-registered to the anatomical image. Lesion maps were manually created based on DWI volumes using MRICron (<http://www.mccauslandcenter.sc.edu/mricron/mricron/>). Then, the time-series were spatially normalized to the standard template of the Montreal Neurological Institute (MNI, Canada)

using the unified segmentation approach with masked lesions [Ashburner and Friston, 2005]. During normalization, all data were resampled to an isotropic voxel size of $1.5 \times 1.5 \times 1.5 \text{ mm}^3$. Finally, data were smoothed with an isotropic Gaussian kernel of 8 mm full-width at half-maximum.

For first-level analysis, we used an auto-regressive model with box-car vectors for each condition (i.e., affected hand movement, unaffected hand movement, instruction) which were convolved with the canonical hemodynamic response function. Head movement parameters were entered as covariates to remove movement-related variance from the time-series [Fox et al., 2005]. A high-pass filter was applied to eliminate signal drifts < 128 s. The individual unthresholded SPM-T-images were used for the multivariate SVM analyses to predict motor outcome at the chronic stage.

Behavioral Assessment

At each session, subjects completed the Action Research Arm Test (ARAT) which is an observer rating of gross and fine upper limb movements on four dimensions including grasp, grip, pinch, and gross movement [Yozbatiran et al., 2008]. The ARAT score for each hand has a range of 0–57. In addition, a grip force index was computed based on the percent maximum grip force of the affected relative to the unaffected hand as measured by a vigorimeter (KLS Martin group, Germany). We computed repeated-measures ANOVAs with the factor “time post-stroke” (i.e., acute, subacute, chronic) for the ARAT score and the grip force index to assess post-stroke recovery.

Definition of Outcome Groups

The idea of the binary support vector machine (SVM) algorithm is to classify new subjects into one of two groups based on a classification rule which has been learned in an independent training data set [Kriegeskorte et al., 2009]. The definition of chronic motor outcome groups was similar to a previous SVM study investigating language outcome in aphasic patients [Saur et al., 2010]. Accordingly, patients were classified into two groups reflecting either “good outcome” or “poor outcome” based on a median split of a *composite motor score* computed from (1) the ARAT scores of the affected hand and (2) the grip force indices at the chronic stage. To obtain these composite motor scores, the individual scores were normalized to a range of 0–1 and then averaged (Table I) [Saur et al., 2010]. Patients with poorer outcome showed a mean ARAT score of 36 (± 26 SD) and a mean grip force index of 42% ($\pm 28\%$ SD). Patients with good outcome featured a mean ARAT score of 56 (± 3 SD) and a mean grip force index of 97% ($\pm 27\%$ SD).

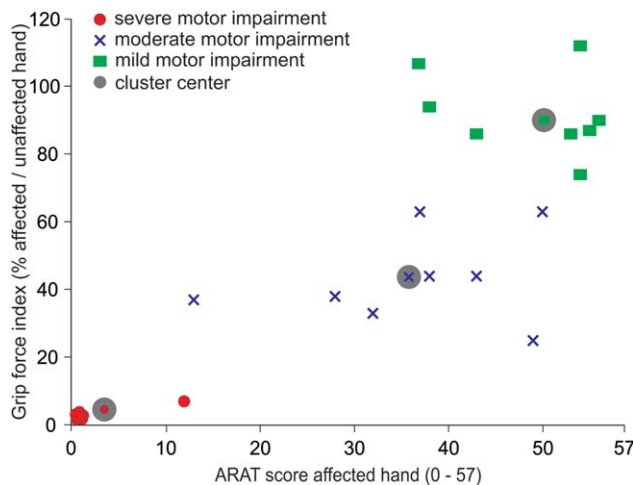


Figure 1.

Initial motor impairment. Results of a *k*-means cluster analysis including both, the ARAT score of the affected hand and the grip force index (in percent relative grip force of the affected relative to the unaffected hand) assessed at the acute stage (6 ± 3 days post-stroke). A three cluster solution provided the best results with at least 12 points distance in the ARAT and 20% difference in grip force between cluster centers. Accordingly, the three clusters were defined as “mild impairment,” “moderate impairment,” and “severe impairment”. Gray shaded symbols indicate the respective cluster centers.

Definition of Initial Motor Impairment

To test whether individual outcome prediction depends on the initial symptom severity [Prabhakaran et al., 2008; Stinear et al., 2012; Zarahn et al., 2011], we used a *k*-means cluster analysis in SPSS to group the ARAT scores and the grip force indices at the acute stage (i.e., 6 ± 3 days post-stroke) into two, three, four, and five clusters. In line with previous studies, the optimal cluster solution was supposed to show at least 12 points distance in the ARAT score and 20% distance in the grip force index [Lang et al., 2008; Stinear et al., 2012]. Only the three cluster solution fulfilled these criteria. The clusters reflected mild, moderate, and severe impairment at the initial assessment (Fig. 1). Patients with mild initial impairment had cluster centers of 50.4 in the ARAT and 90% grip force. Patients with moderate initial impairment showed cluster centers of 36.3 in the ARAT and 43% grip force. Cluster centers in patients with severe initial impairment were 2.0 in the ARAT and 1% grip force.

SVM Features

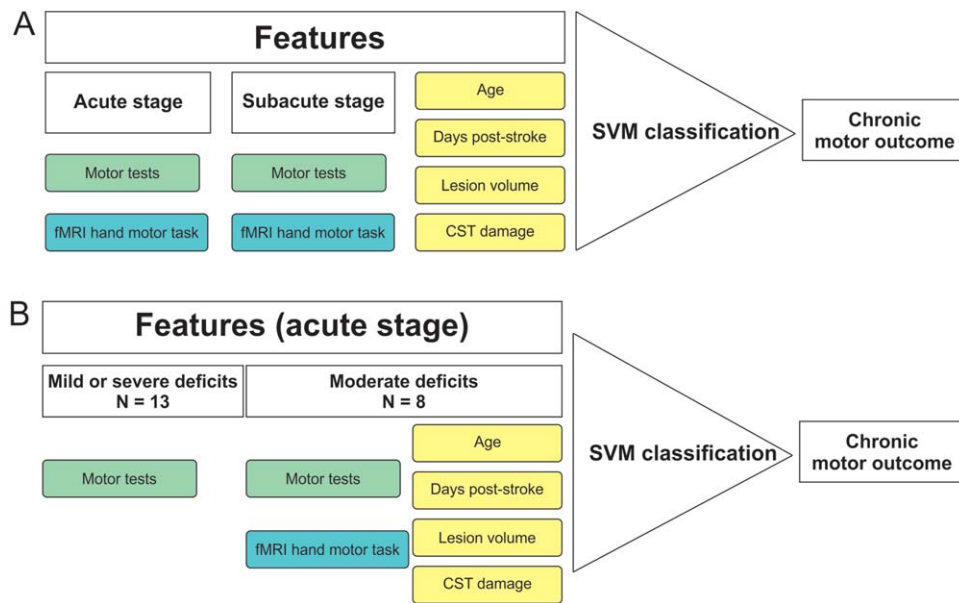
We compared the prognostic accuracy of the following features as potential predictors of motor outcome [Cramer et al., 2007; Cramer, 2008; Stinear, 2010] (Fig. 2): (1) *Assessment of motor function*, including ARAT scores of the affected hand and grip force indices at the acute

and subacute stage, (2) *fMRI features*, consisting of individual unthresholded SPM-T-images for affected hand movements vs. baseline at the acute and subacute stage that were masked by a gray matter mask as defined by the Automated Anatomical Labeling (AAL) atlas to ensure an identical number of voxels per subject [Tzourio-Mazoyer et al., 2002], and (3) *other clinical parameters* including age, time since stroke (from stroke onset to the chronic stage), lesion volume (in mm^3), and percent CST damage which was calculated based on the intersection volume between the MNI-normalized DWI and the myeloarchitectonic maximum probability map of the CST as provided by the Anatomy Toolbox for SPM [Eickhoff et al., 2005]. Prior to the SVM analysis, all features were normalized to range between 0 and 1. We computed separate SVM analyses for each feature as well as SVM analyses for feature combinations of the motor tests, fMRI data, and other clinical parameters. In addition, we compared the classification accuracies of assessments of motor function and fMRI at the acute and subacute stage using McNemar’s chi-squared test of the null hypothesis that classification is equal across these features ($P < 0.05$, one-tailed) [McNemar, 1947]. Power analyses for non-significant differences were calculated using G-Power [Faul et al., 2009].

SVM Analysis

As prediction accuracy may depend on the initial symptom severity, our SVM analyses were twofold. First, single features and feature combinations of the assessment of motor function and fMRI at the acute and subacute stage as well as other clinical parameters from the whole sample of stroke patients ($n = 21$) were used to predict poor vs. good motor outcome at the chronic stage (Fig. 2A). Second, we computed subgroup analyses (1) for patients with either mild or severe initial motor impairment, and (2) patients with moderate initial motor impairment at the acute stage who have been reported to show greater inter-individual variability with regard to good versus poor motor outcome (Fig. 1). We were particularly interested whether the accuracy of prognosis can be improved by fMRI or other clinical data compared to the assessment of motor function in those patients with moderate acute impairment (Fig. 2B).

We used SVM as a supervised machine learning algorithm for the binary classification of motor outcome. Therefore, we computed binary linear SVMs using the LibSVM Library (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>) implemented in Matlab R2014a (TheMathworks, Natick, MA, USA) [Chang and Lin, 2011]. The SVM consisted of a leave-one-subject-out cross-validation algorithm to find a linear decision boundary that optimally separates patients with good vs. poor outcome. Thus, there were n cross-validation loops corresponding to the number of subjects included in the analysis. In each loop, one subject was left out before training the algorithm to optimize the decision

**Figure 2.**

Study design: Prediction of good vs. poor chronic motor outcome. **A:** Features used for SVM classification in a first SVM analysis for the entire sample of stroke patients ($n = 21$). The features comprised the assessment of motor performance (i.e., motor tests including ARAT and grip force) and motor-related fMRI data obtained at the acute and subacute stage post-stroke as well as clinical data including age, days post-stroke, lesion volume, and percent CST damage. **B:** Features used in a second

SVM analysis focusing on three clusters of motor impairment at the acute stage as defined by a cluster analysis (cf. Methods). Patients with initially either mild or severe motor impairment were classified based on the motor performance. Patients with initially moderate impairment show greater outcome variability. Therefore, different features including motor tests, fMRI data, and clinical data were entered into the SVM classification.

boundary for the training data set, i.e. the rest of the sample ($n - 1$) [Rehme et al., in press]. This means that the subjects in the training set and the left-out subject changed in every loop. In each training set, features were selected according to an independent sample t -test ($P < 0.001$). To optimize the decision boundary in the training set in each loop, we varied the soft margin constants (i.e., C -parameter) of the linear separation boundary ranging from small ($C = 0.001$) to large ($C = 30$) in an inner cross-validation loop within the training data set. The model with the highest accuracy and generalizability in the training set was subsequently used to classify the left-out subject based on the features selected in the respective training sample. To control for the danger of circular analysis or “double dipping” [Kriegeskorte et al., 2009], the motor outcome of the left-out subject is always classified based on the decision boundary and the features that have been computed in the training set. This procedure guarantees a naïve prediction of the left-out subject in every loop because training and prediction are completely independent steps. The accuracy of prediction of the left-out subjects across all loops was finally used to determine the mean balanced accuracy of prediction [Brodersen et al., 2010]. Furthermore, we reconstructed the linear SVM weights showing the contribution of different features to the group classification in

each training data set. For voxelwise fMRI data, we approximated the fMRI activation patterns for classification based on the SVM feature weights by calculating the covariance between training data and weight vectors [Haufe et al., 2014]. Finally, we averaged the approximated activation patterns across training samples to get fMRI markers of local activity that predict motor outcome.

For subgroup analysis of initial symptom severity (Fig. 2B), each patient of the two subsamples with either moderate ($n = 8$) or mild and severe initial impairment ($n = 13$) was left out once and predicted based on the model which was optimized for the rest of the entire sample ($n - 1$) as a training set. This resulted in a lower number of cross-validation loops than in the main analysis for all 21 patients. Accordingly, the balanced accuracy and the activation patterns derived from the SVM were averaged across the lower number of loops corresponding to the size of the respective subsample.

RESULTS

Sample

Twenty-one patients who suffered from unilateral hand motor deficits due to first-ever stroke participated in three

TABLE II. SVM classification accuracy for stroke patients with poor ($n = 10$) versus good chronic motor outcome ($n = 11$)

Feature	Mean balanced accuracy (%)	Accuracy poor outcome (%)	Accuracy good outcome (%)	Chi-square, P -value
Single features				
Assessment of motor performance (acute stage)	76%	60%	91%	$\chi^2(1) = 5.762, P = 0.016$
Assessment of motor performance (subacute stage)	75%	50%	100%	$\chi^2(1) = 5.762, P = 0.016$
fMRI (acute stage)	86%	90%	82%	$\chi^2(1) = 10.714, P = 0.001$
fMRI (subacute stage)	66%	—	—	$\chi^2(1) = 2.33, P = 0.127$ n. s.
Other clinical data (Age, time since stroke, lesion volume, % CST damage)	all $\leq 50\%$	—	—	$\chi^2(1) = 0.048, P = 0.827$ n. s.
Feature combinations				
fMRI + motor performance (acute)	86%	90%	82%	$\chi^2(1) = 10.714, P = 0.001$
fMRI (acute) + clinical data	86%	90%	82%	$\chi^2(1) = 10.714, P = 0.001$
Motor performance (acute) + clinical data	71%	—	—	$\chi^2(1) = 3.857, P = 0.05$ n. s.
fMRI + motor performance (acute) + clinical data	86%	90%	82%	$\chi^2(1) = 10.714, P = 0.001$
fMRI + motor performance (subacute)	61%	—	—	$\chi^2(1) = 0.429, P = 0.513$ n. s.
fMRI (subacute) + clinical data	71%	—	—	$\chi^2(1) = 3.857, P = 0.05$ n. s.
Motor performance (subacute) + clinical data	75%	50%	100%	$\chi^2(1) = 5.762, P = 0.016$
fMRI + motor performance (subacute stage) + clinical data	66%	—	—	$\chi^2(1) = 2.33, P = 0.127$ n. s.

Assessment of motor performance refers to the ARAT score of the affected hand and the grip force index of the affected relative to the unaffected hand (in %). Other clinical data refers to features including age, days post-stroke (at the final follow-up assessment at the chronic stage), lesion volume, and percent corticospinal tract damage. ARAT: Action Research Arm Test (score for the affected hand, maximum = 57); CST: corticospinal tract; n.s.: not significant.

sessions at the acute, subacute, and chronic stage post-stroke (Table I). The maximum lesion overlap was along the CST (Supporting Information Figs. 1 and 2). The degree of impairment varied from hemiplegia to mild hand weakness—a typical range of deficits encountered in a clinical sample. Patients with good outcome at the chronic stage showed less initial impairment at the acute and subacute stage (i.e., higher ARAT scores and grip force indices) than patients with relatively poor outcome (all t -tests $P \leq 0.021$). Nevertheless, all except of three patients showed substantial improvements in hand motor function from the acute to the chronic stage. A repeated-measures ANOVA revealed a significant main effect of “time post-stroke” in the ARAT ($F(2,40) = 14.719, P < 0.001$) and the grip force index ($F(2,40) = 6.188, P = 0.005$).

Prediction of Chronic Motor Outcome After Stroke

The first analysis was used to predict motor outcome at the chronic stage for the entire sample of stroke patients ($n = 21$) based on the assessment of motor performance,

fMRI, and other clinical data obtained at the acute and subacute stage post-stroke (Fig. 2A). The level of motor impairment (i.e., ARAT scores and grip force indices) at the acute stage predicted chronic motor outcome with a mean balanced classification accuracy of 75.5% ($\chi^2(1) = 5.762, P = 0.016$) (Table II). While both motor tests provided a high accuracy of 91% to classify patients with good outcome, the accuracy to correctly predict patients with poor outcome was low (i.e., 60%). The prediction of motor outcome based on the motor performance assessed at the subacute stage was similar to that of the acute stage (mean accuracy: 75%, $\chi^2(1) = 5.762, P = 0.016$; sensitivity of correctly classifying poor outcome: 50%, specificity: 100%). Taken together, the motor tests had a low prognostic value to identify those patients with poor motor outcome.

In contrast, the SVM analysis based on task-related fMRI activity for voluntary movements of the affected hand at the acute stage classified good versus poor motor outcome at the chronic stage with 86% accuracy ($\chi^2(1) = 10.714, P = 0.001$). fMRI data provided high accuracy of 90% to identify patients with poor outcome (i.e., high sensitivity) and 82% accuracy to predict patients with

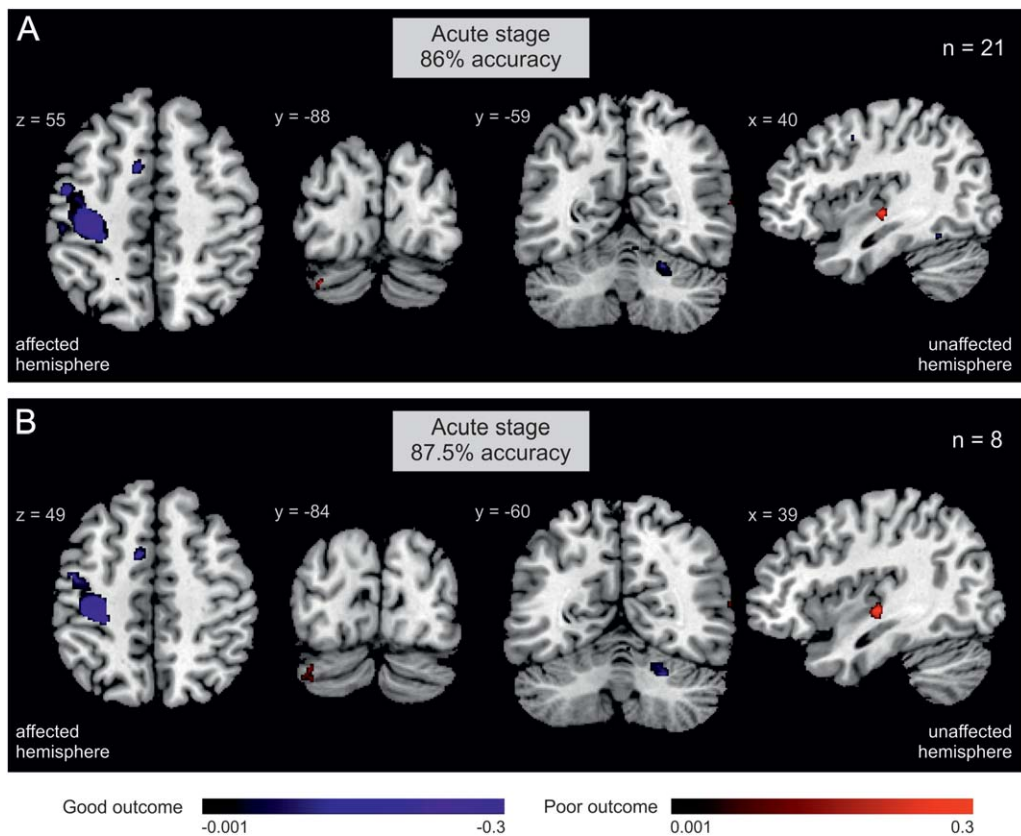


Figure 3.

Activation patterns derived from the SVM that show regional activity contributing to outcome classification. **A:** Regional activity and prediction accuracy of chronic motor outcome based on motor-related fMRI data acquired at the acute stage in the whole sample ($n = 21$). There was no significant outcome classification based on subacute fMRI data. **B:** Regional activity and prediction accuracy of chronic motor outcome based on motor-

related fMRI data acquired at the acute stage for stroke patients with moderate impairment at the acute stage ($n = 8$). Blue = regional activity contributing to the classification of good motor outcome (i.e., patients with poor outcome show less activity in these areas); red = regional activity classifying poor motor outcome.

good outcome (i.e., high specificity). Importantly, there was no difference in accuracy for the two experimental samples ($\chi^2(1) = 2.27$, $P = 0.264$), showing that differences in the fMRI sequence did not affect the SVM results. Likewise, there was no difference in the location of stroke lesions between correctly and incorrectly classified patients (Supporting Information Figs. 2 and 3).

When mapping voxels that contributed to the outcome classification, we found that patients with higher levels of activity in ipsilesional primary motor cortex (M1), supplementary motor area (SMA), dorsal premotor cortex, and contralesional cerebellum (lobule V, VI) at the acute stage featured good motor outcome (Fig. 3A). In turn, patients with poor motor outcome featured reduced activity in these areas. Furthermore, patients with poor outcome were likely to show enhanced activity in contralesional mid-posterior insula and ipsilesional cerebellum (crus I) at the acute stage. In contrast to acute fMRI, motor-related fMRI activity

obtained at the subacute stage provided no significant classification of chronic outcome (mean accuracy: 66%; $\chi^2(1) = 2.33$, $P = 0.127$). The difference in prediction accuracy between acute and subacute fMRI data showed a trend toward significance ($\chi^2(1) = 2.94$, $P = 0.09$). The classification accuracies of other clinical parameters including age, time since stroke, lesion volume, and percent CST damage were below 50% chance-level (Table II).

Different combinations of motor tests, fMRI, and other clinical parameters did not exceed the classification accuracy of the SVM for acute stage fMRI alone (please see Table II for detailed results).

Initial Impairment and Chronic Motor Outcome

The first analysis of the whole sample showed that the level of motor impairment was a rather weak predictor especially for patients with poor motor outcome.

TABLE III. SVM classification accuracy for poor ($n = 4$) versus good motor outcome ($n = 4$) in a subsample of stroke patients with moderate initial motor impairment ($n = 8$)

Feature	Mean balanced accuracy (%)	Sensitivity (%)	Specificity (%)	Chi-square, P -value
Single features				
Assessment of motor performance (acute)	63%	—	—	$\chi^2(1) = 0.5, P = 0.480$ n. s.
fMRI (acute)	87.5%	100%	75%	$\chi^2(1) = 4.5, P = 0.034$
Clinical data (age, time since stroke, lesion volume, % CST damage)	$\leq 50\%$	—	—	$\chi^2(1) = 0.143, P = 0.705$ n.s.
Feature combinations				
fMRI + motor performance (acute)	87.5%	100%	75%	$\chi^2(1) = 4.5, P = 0.034$
fMRI (acute) + clinical data	87.5%	100%	75%	$\chi^2(1) = 4.5, P = 0.034$
Motor performance (acute) + clinical data	50%	—	—	$\chi^2(1) = 0.143, P = 0.705$ n.s.
fMRI + motor performance (acute) + clinical data	87.5%	100%	75%	$\chi^2(1) = 4.5, P = 0.034$

Assessment of motor performance refers to the ARAT score of the affected hand and the grip force index of the affected relative to the unaffected hand (in %). ARAT: Action Research Arm Test (score for the affected hand, maximum = 57); CST: corticospinal tract; n.s.: not significant.

Therefore, we tested the hypothesis that the prognostic value of different features (i.e., fMRI, assessments of motor function, clinical parameters) depends on the degree of initial motor impairment in a subgroup analysis considering initial symptom severity (Fig. 2B, cf. Methods). To this end, we performed different SVM analyses (i) for patients with mild and severe initial impairment, and (ii) for patients with moderate initial impairment.

Entering the motor test scores (i.e., ARAT, grip force index) from the acute stage into the SVM algorithm provided a significant outcome classification for the subsample with either *mild* or *severe* motor deficits at the acute stage with 85% accuracy ($\chi^2(1) = 6.231, P = 0.013$) (Table III), but not for the subgroup of patients with initially *moderate* deficits (62.5% accuracy; $\chi^2(1) = 0.5, P = 0.480$). In contrast, motor-related fMRI activity obtained at the acute stage yielded 87.5% classification accuracy in patients with moderate initial impairment ($\chi^2(1) = 4.5, P = 0.034$). The regions contributing to the prediction of good vs. poor outcome in initially moderately impaired patients were similar to those observed for the outcome prediction of the entire sample (Fig. 3B): Activity in ipsilesional M1, SMA, and contralesional cerebellum predicted good motor outcome whereas additional activity in contralesional insula and ipsilesional cerebellum was a marker for poor outcome. Again, the accuracy of other features or combinations of features did not lead to higher classification accuracies than fMRI data alone (please see Table III for details). Although assessments of motor function did not provide a significant prediction of motor outcome in contrast to the acute fMRI data, the difference of these two predictions was not significant ($\chi^2(1) = 1, P = 0.309$). An

additional power analysis showed that a sample size of 53 patients would have been required to find a statistically significant difference with 80% probability.

In summary, early motor tests failed to reliably predict chronic motor outcome in the subsample with moderate initial impairment: Half of these patients substantially recovered whereas the other half suffered from poor motor outcome (Table I). In contrast, acute fMRI data yielded a significant prediction in the subsample with moderate initial impairment

DISCUSSION

We found that fMRI data of the active motor system obtained at the acute stage post-stroke predict chronic motor outcome with high prognostic accuracy. By contrast, outcome prediction by motor impairment was less robust. Notably, the assessment of motor function was sufficient to predict the outcome of patients with either mild or severe initial deficits. In contrast, patients with moderate initial impairment and greater variability of motor outcome could not be classified based on motor tests alone. In these patients, fMRI of the motor system provides additional information that enables a more accurate prediction of the individual motor outcome.

The Prognostic Use of fMRI in Post-Stroke Recovery

The data of the present study suggest that the use of SVMs for analyzing fMRI data has the potential to serve

as a diagnostic marker for post-stroke outcome [Cramer et al., 2007; Stinear and Ward, 2013]. In line with findings from longitudinal studies in humans and animals showing that the degree of a chronic deficit is mainly defined within the first days post-stroke, classification accuracy was only high for motor-related fMRI data obtained at the acute stage (6 ± 3 days) after stroke [Carmichael, 2003; Cramer, 2008; Kwakkel and Kollen, 2007]. In contrast, fMRI acquired after the second week post-stroke (12 ± 4 days) provided no significant outcome classification. Accordingly, there seems to be a sensitive time window of fMRI in the first few days after stroke for later outcome prediction. This time dependency is likely to result from a close relationship between motor impairment and loss of fMRI activity in the acute stage post-stroke [Rehme et al., in press]. In contrast, subsequent cortical reorganization in the subacute stage may change the motor-related fMRI activity pattern in a way that also patients with more severe deficits feature high levels of fMRI activity [Rehme et al., 2011b; Ward et al., 2003]. Importantly, the prognostic value of fMRI motor tasks does not depend on particular fMRI parameters as our data show no difference in accuracy between two MRI sequences. Thus, statistical maps derived from task-related fMRI seem to be robust against technical differences as long as the experimental design induces sufficient motor-related activity which is usually the case for block designs lasting at least 5 min [Rehme et al., 2011b]. This may facilitate the implementation of fMRI into standard clinical protocols.

Previous fMRI group studies provided evidence that fMRI activity for movements of the stroke-affected hand correlates with motor outcome, and therefore may contain relevant information for prognosis at the level of individual patients [Cramer et al., 2007; Marshall et al., 2009; Zarahn et al., 2011]. As these studies were based on regression analyses, inference was restricted to the group-level. However, evidence-based disease models require the identification of prognostic markers at the level of individual patients [Orrù et al., 2012]. New technological advances allow using fMRI in the framework of multivariate machine learning to decode individual markers from brain activation patterns [Naselaris et al., 2011]. We recently demonstrated that resting-state fMRI allows classification of individual motor impairment in acute stroke patients using SVM. This finding provides a proof-of-principle that multivariate decoding of fMRI data may be used to infer the patients' current functional state [Rehme et al., in press]. Notably, the present study extends previous work showing that fMRI may even be used to predict the future functional outcome for individual patients. Similar to our findings in patients with motor disability, Saur et al. reported that SVM predicts language outcome in stroke patients with aphasia based on language-related fMRI activity [Saur et al., 2010]. The high prognostic accuracy of fMRI in contrast to assessments of motor function is likely based on the multidimensional nature of voxelwise brain

activity. Technically, multivariate SVM pattern classification based on multidimensional fMRI data has greater power and is less prone to outliers than simple motor tests because the SVM considers voxelwise dependencies to better separate different outcome classes. In contrast, motor tests consist of fewer dimensions to map relevant features of motor outcome. Furthermore, fMRI may provide insights into biologically relevant phenotypes of brain recovery. Motor function is mediated by different neural processes that can be supporting or detrimental for brain reorganization after stroke. For example, the activity of contralesional M1 has been shown to support motor function at the early stage after stroke [Rehme et al., 2011a], but may turn into a maladaptive inhibitory influence toward the affected hemisphere that maintains motor disability at the chronic stage [Grefkes et al., 2008]. Together, the findings of the present study in line with previous evidence [Rehme et al., in press; Saur et al. 2010] suggest that fMRI has a relevant diagnostic potential that may be used to predict the outcome of neurological deficits.

Predictors of Chronic Impairment

Thus far, only a few factors have been identified which reliably predict the long-term outcome after stroke [van Almenkerk et al., 2013]. Stroke severity is one important factor as patients with more severe neurological deficits more often remain permanently disabled [Kwakkel and Kollen, 2013]. Consistently, we also found that patients with initially more severe deficits had a less favorable outcome. In contrast, patients with moderate impairment are much more variable with respect to their outcome, a finding which is confirmed by the present study and which limits the prognostic value of behavioral parameters alone [Prabhakaran et al., 2008; Stinear, 2010; Stinear et al., 2012; Zarahn et al., 2011].

Likewise, lesion volume is an unreliable marker for recovery. Massive lesions affecting the entire MCA territory are usually associated with poor outcome, but the relationship between smaller lesions and outcome is more variable [Rehme et al., in press; Schiemanck et al., 2006]. In line with these data, we did not find a predictive value of lesion size in our sample. The factor "age" has been suggested repeatedly to predict better outcome especially for young patients [Maaijwee et al., in press]. Saur et al. reported that age improves individual SVM prognosis of fMRI data in aphasic stroke patients [Saur et al., 2010]. Compared to the study by Saur et al., our sample did not include a sufficient number of younger patients which may explain that age did not influence the classification of motor outcome in the present study.

Neuroimaging Markers of Outcome

fMRI obtained in the first week post-stroke provided an accurate prediction of motor outcome (i.e., 86%).

Classification was mainly driven by activity in ipsilesional M1, premotor areas, and contralesional cerebellum during movements of the affected hand (Fig. 3). M1 is the primary origin of corticospinal neurons controlling motor execution [Dum and Strick, 2002]. Neurons in premotor areas including premotor cortex and SMA are involved in movement preparation via dense projections to M1 [Dum and Strick, 2002; Rouiller et al., 1994]. Fine-tuning of movements is further mediated by cortico-cerebellar loops terminating in lobules IV–VI of anterior cerebellum [Middleton and Strick, 1994; Stoodley and Schmahmann, 2009]. In line with our findings, previous fMRI group studies showed that the increase of activity in ipsilesional M1 is associated with spontaneous motor improvements, better motor outcome, and training effects [Cramer et al., 2007; Loubinoux et al., 2007; Rehme et al., 2011b, 2012; Tombari et al., 2004]. Likewise, enhanced activity in ipsilesional SMA has been shown to positively influence ipsilesional M1 activity after stroke which drives functional recovery [Loubinoux et al., 2007; Rehme et al., 2011a; Rehme and Grefkes, 2013]. Activity in contralesional cerebellum has also been reported to positively correlate with motor performance and outcome [Loubinoux et al., 2007; Rehme et al., 2012; Small et al., 2002]. In contrast, fMRI studies found that activity in ipsilesional cerebellum is enhanced in stroke patients with greater motor deficits relative to controls but decreases as motor functions recover [Rehme et al., 2012; Ward et al., 2003]. In our study, activation of the posterior part (i.e., crus I in lobule VIIA) of ipsilesional cerebellum at the acute stage decodes poor motor outcome. Whereas anterior cerebellum is involved in sensorimotor processing, the posterior part has been suggested to mediate cognitive tasks via loops with higher-order association areas in posterior-parietal and prefrontal cortex [Brodal and Bjaalie, 1997; Stoodley and Schmahmann, 2009]. Thus, the observed activation of ipsilesional cerebellum during movements of the affected hand may reflect increased recruitment of higher-order executive functions in an effort to initiate movement and to compensate for disrupted primary sensorimotor pathways. Patients with poor outcome were also characterized by greater activity in contralesional mid-posterior insula. This part of the insula has been demonstrated to integrate sensory and motor-related information from primary and secondary sensorimotor areas [Kurth et al., 2010]. One hypothesis is that the greater recruitment of contralesional insula may reflect enhanced processing of somatosensory feedback to support residual motor function.

Motor Outcome and CST Damage

CST damage has been shown to be a strong correlate of acute motor deficits and recovery thereof [Stinear and Ward, 2013]. For example, studies using transcranial mag-

netic stimulation (TMS) demonstrated that reduced excitability of the CST is related to greater motor impairment and less recovery [Stinear et al., 2007, 2012; Swayne et al., 2008]. Likewise, diffusion tensor imaging (DTI) studies provided evidence that the greater the disruption of the structural integrity of the CST, particularly at the posterior limb of the internal capsule, the more limited is functional recovery [Stinear and Ward, 2013]. Recently, Stinear et al. presented an empirically validated decision algorithm that allows predicting motor outcome based on the serial combination of motor tests followed by TMS and DTI measures [Stinear et al., 2012]. In line with our findings, the authors reported that accurate prognosis of patients with more severe impairment requires additional information about the individual network pathology. Importantly, however, and in contrast to their results, we did not find any significant difference in CST damage between our two outcome groups (Supporting Information Fig. 1). These discrepant findings probably result from different ways of assessing CST integrity. While we estimated damage of the entire CST based on myeloarchitectonic probability maps, Stinear et al. used DTI to focus on the fiber tract integrity at the level of the internal capsule. Given the high density of fibers at this level, it is well conceivable that patients with stronger damage to this region suffer from poorer recovery. Notably, there is a strong correlation between CST damage—both globally and at the level of the internal capsule—and motor-related fMRI activity [Hendricks et al., 2002; Stinear et al., 2007; Swayne et al., 2008]. Thus, the fMRI signal likely reflects a combination of structural and functional brain states [Ward et al., 2006]. Our data, therefore, suggest that fMRI provides a composite measure of network pathology following stroke which captures the effect of CST disruptions. Future studies should test the predictive power of SVMs for stroke recovery based on DTI parameters in combination with fMRI.

CONCLUSION

Motor-related fMRI data provided good prognostic accuracy at the level of individual patients in contrast to behavioral assessments of motor function that were less reliable. This was particularly evident in initially moderately impaired patients who demonstrated high variability in motor recovery [Prabhakaran et al., 2008; Stinear, 2010; Stinear et al., 2012; Zarahm et al., 2011]. Although there was no significant difference between predictions based on fMRI or assessments of motor function because of small sample sizes, our findings show that fMRI in the first few days after stroke can add substantial information to predictive models independent of baseline clinical impairment [Cramer et al., 2007; Saur et al., 2010]. This finding may motivate future trials with larger samples to validate the multivariate machine learning approach in an independent cohort. Furthermore, the implementation of fMRI in decision algorithms may establish the added

predictive value of fMRI in clinical practice [Orrù et al., 2012; Stinear et al., 2012]. Such studies will pave the way toward evidence-based clinical decision-making and more accurate stratification of patients into clinical trials [Stinear and Ward, 2013].

REFERENCES

- Ashburner J, Friston KJ (2005): Unified segmentation. *Neuroimage* 26:839–851.
- Brodal P, Bjaalie JG (1997): Salient anatomic features of the cortico-ponto-cerebellar pathway. *Prog Brain Res* 114:227–249.
- Brodersen KH, Ong CS, Stephan KE, Buhmann JM (2010): The balanced accuracy and its posterior distribution. *Int Conf Pattern Recogn* 3121–3124.
- Carminichael ST (2003): Plasticity of cortical projections after stroke. *Neuroscientist* 9:64–75.
- Chang CC, Lin CJ (2011): LIBSVM: A library for support vector machines. *ACM Trans Intell Syst Technol* 27:1–39.
- Cramer SC (2008): Repairing the human brain after stroke. I. Mechanisms of spontaneous recovery. *Ann Neurol* 63:272–287.
- Cramer SC, Parrish TB, Levy RM, Stebbins GT, Ruland SD, Lowry DW, Trouard TP, Squire SW, Weinand ME, Savage CR, Wilkinson SB, Juranek J, Leu SY, Himes DM (2007): Predicting functional gains in a stroke trial. *Stroke* 38:2108–2114.
- Dum RP, Strick PL (2002): Motor areas in the frontal lobe of the primate. *Physiol Behav* 77:677–682.
- 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.
- Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005): A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25:1325–1335.
- Faul F, Erdfelder E, Buchner A, Lang A-G (2009): Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav Res Methods* 41:1149–1160.
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005): The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci USA* 102:9673–9678.
- Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Borden WB, Bravata DM, Dai S, Ford ES, Fox CS, Franco S, Fullerton HJ, Gillespie C, Hailpern SM, Heit JA, Howard VJ, Huffman MD, Kissela BM, Kittner SJ, Lackland DT, Lichtman JH, Lisabeth LD, Magid D, Marcus GM, Marelli A, Matchar DB, McGuire DK, Mohler ER, Moy CS, Mussolino ME, Nichol G, Paynter NP, Schreiner PJ, Sorlie PD, Stein J, Turan TN, Virani SS, Wong ND, Woo D, Turner MB; American Heart Association Statistics Committee and Stroke Statistics Subcommittee (2013): Heart disease and stroke statistics–2013 update: A report from the American Heart Association. *Circulation* 127:e6–e245.
- Grefkes C, Eickhoff SB, Nowak DA, Dafotakis M, Fink GR (2008): Dynamic intra- and interhemispheric interactions during unilateral and bilateral hand movements assessed with fMRI and DCM. *Neuroimage* 41:1382–1394.
- Haufe S, Meinecke F, Gorgen K, Dahne S, Haynes JD, Blankertz B, Bießmann F (2014): On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage* 87:96–110.
- Hendricks HT, Zwartz MJ, Plat EF, van LJ (2002): Systematic review for the early prediction of motor and functional outcome after stroke by using motor-evoked potentials. *Arch Phys Med Rehabil* 83:1303–1308.
- Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009): Circular analysis in systems neuroscience: The dangers of double dipping. *Nat Neurosci* 12:535–540.
- Kurth F, Zilles K, Fox PT, Laird AR, Eickhoff SB (2010): A link between the systems: Functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Struct Funct* 214:519–534.
- Kwakkel G, Kollen B (2007): Predicting improvement in the upper paretic limb after stroke: A longitudinal prospective study. *Restor Neurol Neurosci* 25:453–460.
- Kwakkel G, Kollen BJ (2013): Predicting activities after stroke: What is clinically relevant? *Int J Stroke* 8:25–32.
- Lang CE, Edwards DF, Birkenmeier RL, Dromerick AW (2008): Estimating minimal clinically important differences of upper-extremity measures early after stroke. *Arch Phys Med Rehabil* 89:1693–1700.
- Loubinoux I, Dechaumont-Palacin S, Castel-Lacanal E, De B, Marque X, Pariente P, Albucher J, Berry JF, Chollet IF (2007): Prognostic value of fMRI in recovery of hand function in sub-cortical stroke patients. *Cereb Cortex* 17:2980–2987.
- Maaijwee NA, Arntz RM, Rutten-Jacobs LC, Schaapsmeeders P, Schoonderwaldt HC, van Dijk EJ, de Leeuw FE: Post-stroke fatigue and its association with poor functional outcome after stroke in young adults. *J Neurol Neurosurg Psychiatry* (in press).
- 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.
- McNemar Q (1947): Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12:153–157.
- Middleton FA, Strick PL (1994): Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science* 266:458–461.
- Miller EL, Murray L, Richards L, Zorowitz RD, Bakas T, Clark P, Billinger SA (2010): Comprehensive overview of nursing and interdisciplinary rehabilitation care of the stroke patient: A scientific statement from the American Heart Association. *Stroke* 41:2402–2448.
- Naselaris T, Kay KN, Nishimoto S, Gallant JL (2011): Encoding and decoding in fMRI. *Neuroimage* 56:400–410.
- 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.
- Orrù G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A (2012): Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neurosci Biobehav Rev* 36:1140–1152.
- 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.
- Rehme AK, Grefkes C (2013): Cerebral network disorders after stroke: Evidence from imaging-based connectivity analyses of active and resting brain states in humans. *J Physiol* 591:17–31.

- Rehme AK, Eickhoff SB, Wang LE, Fink GR, Grefkes C (2011a): Dynamic causal modeling of cortical activity from the acute to the chronic stage after stroke. *Neuroimage* 55:1147–1158.
- Rehme AK, Fink GR, von Cramon DY, Grefkes C (2011b): The role of the contralesional motor cortex for motor recovery in the early days after stroke assessed with longitudinal fMRI. *Cereb Cortex* 21:756–768.
- Rehme AK, Eickhoff SB, Rottschy C, Fink GR, Grefkes C (2012): Activation likelihood estimation meta-analysis of motor-related neural activity after stroke. *Neuroimage* 59:2771–2782.
- Rehme AK, Volz LJ, Feis DL, Bomilcar-Focke I, Liebig T, Eickhoff SB, Fink GR, Grefkes C: Identifying neuroimaging markers of motor disability in acute stroke by machine learning techniques. *Cereb Cortex* (in press).
- Rouiller EM, Babalian A, Kazennikov O, Moret V, Yu XH, Wiesendanger M (1994): Transcallosal connections of the distal forelimb representations of the primary and supplementary motor cortical areas in macaque monkeys. *Exp Brain Res* 102: 227–243.
- Saur D, Ronneberger O, Kummerer D, Mader I, Weiller C, Kloppel S (2010): Early functional magnetic resonance imaging activations predict language outcome after stroke. *Brain* 133: 1252–1264.
- Schiemanck SK, Kwakkel G, Post MW, Prevo AJ (2006): Predictive value of ischemic lesion volume assessed with magnetic resonance imaging for neurological deficits and functional outcome poststroke: A critical review of the literature. *Neurorehabil Neural Repair* 20:492–502.
- Small SL, Hlustik P, Noll DC, Genovese C, Solodkin A (2002): Cerebellar hemispheric activation ipsilateral to the paretic hand correlates with functional recovery after stroke. *Brain* 125:1544–1557.
- Stinear C (2010): Prediction of recovery of motor function after stroke. *Lancet Neurol* 9:1228–1232.
- Stinear CM, Ward NS (2013): How useful is imaging in predicting outcomes in stroke rehabilitation? *Int J Stroke* 8:33–37.
- Stinear CM, Barber PA, Smale PR, Coxon JP, Fleming MK, Byblow WD (2007): Functional potential in chronic stroke patients depends on corticospinal tract integrity. *Brain* 130: 170–180.
- Stinear CM, Barber PA, Petoe M, Anwar S, Byblow WD (2012): The PREP algorithm predicts potential for upper limb recovery after stroke. *Brain* 135:2527–2535.
- Stoodley CJ, Schmahmann JD (2009): Functional topography in the human cerebellum: A meta-analysis of neuroimaging studies. *Neuroimage* 44:489–501.
- 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.
- Tombari D, Loubinoux I, Pariente J, Gerdelat A, Albucher JF, Tardy J, Cassol E, Chollet F (2004): A longitudinal fMRI study: In recovering and then in clinically stable sub-cortical stroke patients. *Neuroimage* 23:827–839.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002): Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15:273–289.
- van Almenkerk S, Smalbrugge M, Depla MF, Eefsting JA, Hertogh CM (2013): What predicts a poor outcome in older stroke survivors? A systematic review of the literature. *Disabil Rehabil* 35:1774–1782.
- Veerbeek JM, Kwakkel G, van Wegen EE, Ket JC, Heymans MW (2011): Early prediction of outcome of activities of daily living after stroke: A systematic review. *Stroke* 42:1482–1488.
- Ward NS, Brown MM, Thompson AJ, Frackowiak RS (2003): Neural correlates of motor recovery after stroke: A longitudinal fMRI study. *Brain* 126:2476–2496.
- Ward NS, Newton JM, Swayne OB, Lee L, Thompson AJ, Greenwood RJ, Rothwell JC, Frackowiak RS (2006): Motor system activation after subcortical stroke depends on corticospinal system integrity. *Brain* 129:809–819.
- Yozbatiran N, Der-Yeghiaian L, Cramer SC (2008): A standardized approach to performing the action research arm test. *Neurorehabil Neural Repair* 22:78–90.
- Zarahn E, Alon L, Ryan SL, Lazar RM, Vry MS, Weiller C, Marshall RS, Krakauer JW (2011): Prediction of motor recovery using initial impairment and fMRI 48 h poststroke. *Cereb Cortex* 21:2712–2721.