The PREP algorithm predicts potential for upper limb recovery after stroke

Cathy M. Stinear,1,2 P. Alan Barber,1,2,3 Matthew Petoe,1,2 Samir Anwar2,4 and Winston D. Byblow2,5

1 Department of Medicine, University of Auckland, Auckland 1142, New Zealand
2 Centre for Brain Research, University of Auckland, Auckland 1142, New Zealand
3 Neurology Department, Auckland City Hospital, Auckland 1023, New Zealand
4 RehabPlus, Auckland District Health Board, Auckland 1025, New Zealand
5 Department of Sport and Exercise Science, University of Auckland, Auckland 1142, New Zealand

Correspondence to: Professor Winston Byblow, Department of Sport and Exercise Science and Centre for Brain Research, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
E-mail: w.byblow@auckland.ac.nz

Stroke is a leading cause of adult disability and the recovery of motor function is important for independence in activities of daily living. Predicting motor recovery after stroke in individual patients is difficult. Accurate prognosis would enable realistic rehabilitation goal-setting and more efficient allocation of resources. The aim of this study was to test and refine an algorithm for predicting the potential for recovery of upper limb function after stroke. Forty participants were prospectively enrolled within 3 days of ischaemic stroke. First, shoulder abduction and finger extension strength were graded 72 h after stroke onset to compute a shoulder abduction and finger extension score. Secondly, transcranial magnetic stimulation was used to assess the functional integrity of descending motor pathways to the affected upper limb. Third, diffusion-weighted magnetic resonance imaging was used to assess the structural integrity of the posterior limbs of the internal capsules. Finally, these measures were combined in the PREP algorithm for predicting an individual’s potential for upper limb recovery at 12 weeks, measured with the Action Research Arm Test. A cluster analysis was used to independently group patients according to Action Research Arm Test score at 12 weeks, for comparison with predictions from the PREP algorithm. There was excellent correspondence between the cluster analysis of Action Research Arm Test score at 12 weeks and predictions made with the PREP algorithm. The algorithm had positive predictive power of 88%, negative predictive power of 83%, specificity of 88% and sensitivity of 73%. This study provides preliminary data in support of the PREP algorithm for the prognosis of upper limb recovery in individual patients. PREP may enable tailored planning of rehabilitation and more accurate stratification of patients in clinical trials.

Keywords: stroke; prognosis; rehabilitation; motor; TMS; MRI

Abbreviations: ARAT = Action Research Arm Test; FA = fractional anisotropy; FM = Fugl-Meyer; MEP = motor-evoked potential; NIHSS = National Institutes of Health Stroke Scale; PREP = predicting recovery potential; SAFE = shoulder abduction and finger extension; TMS = transcranial magnetic stimulation
Introduction

Stroke is the second most common cause of death worldwide and the third most common cause of long-term adult disability in developed countries (WHO, 2004). Significant advances have been made in the diagnosis, prevention and treatment of stroke in recent decades (Hachinski et al. 2010), but advances in stroke rehabilitation have been more modest. Independence in activities of daily living depends to a great extent on the recovery of motor function, particularly in the upper limb (Veerbeek et al. 2011). The estimated recovery of motor function is one of several factors that influence decisions regarding the goals, type and duration of rehabilitation for each patient. These decisions have far-reaching consequences, but are usually made in the absence of objective information regarding the patient’s capacity for motor recovery. Accurate prognosis of an individual’s potential for recovery would enable realistic goal-setting, guide the allocation of rehabilitation resources and help to manage expectations.

Clinical measures of upper limb impairment made within days of stroke are related to subsequent outcome (Katradis et al. 1998; Smania et al. 2007; Nijland et al. 2010; Veerbeek et al. 2011), but have little individual prognostic value due to interindividual variability (Stinear 2010). Neurophysiology and neuroimaging techniques may also have a role in determining recovery potential. Transcranial magnetic stimulation (TMS) can be used to test the functional integrity of corticospinal pathways originating from the ipsilesional primary motor cortex (M1). In general, a patient has better prospects for upper limb recovery if TMS can elicit motor-evoked potentials (MEPs) in affected upper limb muscles within days of stroke (Rapisarda et al. 1996; Escudero et al. 1998; Cruz Martinez et al. 1999). However, some individuals with no upper limb MEPs may still recover some manual dexterity (Heald et al. 1993). Diffusion-weighted MRI can be used to measure the structural integrity of descending white matter pathways. Studies using diffusion-weighted imaging have shown that greater loss of ipsilesional tract integrity (Schiemanck et al. 2008; Radlinska et al. 2010), at the level of the posterior limb of the internal capsule (Shelton and Reding, 2001; Schiemanck et al. 2008; Puig et al., 2011) or brainstem (DeVetten et al. 2010) is associated with worse motor outcome. Most previous studies have used clinical, TMS or MRI measures in isolation and have correlated these to motor outcomes in groups and not individual patients. While these studies have improved our understanding of stroke recovery, a shift away from correlational approaches is needed before these measures can be used to guide individual rehabilitation decisions (Ward 2011).

We previously found that a combination of TMS and MRI measures could predict the potential for further reductions in upper limb impairment at the chronic stage of stroke recovery (Stinear et al. 2007). Based on these data, an algorithm was developed for stratifying chronic stroke patients and selecting rehabilitation strategies accordingly. We then proposed a new algorithm that could be used in patients with acute and sub-acute stroke (Fig. 1; Stinear 2010): the predicting recovery potential (PREP) algorithm for the upper limb. The main difference between the PREP algorithm for acute patients and its predecessor for chronic patients is that it begins with an assessment of shoulder abduction and finger extension impairment in the affected upper limb at 72 h after symptom onset. These movements are graded using the Medical Research Council scoring system, and the two scores are added to give a SAFE score (Shoulder Abduction Finger Extension, range 0–10). This clinical component was included as the presence or absence of these movements at 72 h has some power to predict upper limb function at 6 months (Nijland et al. 2010).

Each element of the PREP algorithm allows the sequential identification of a subset of patients. A SAFE score of ≥ 8 allows the separation of patients with predicted ‘complete’ recovery from three other predicted outcome categories (notable, limited or none; Table 1). Only patients with a SAFE score <8 proceed to neurophysiological and possibly neuroimaging assessments. The presence of upper limb MEPS in response to TMS allows the separation of those with predicted ‘notable’ recovery from the remaining two predicted outcome categories. Diffusion-weighted imaging can then be used to separate patients in these final two categories into those with ‘limited’ or ‘none’ recovery potential. The PREP algorithm is predicated on the evidence that sparing of descending white matter pathways is related to better recovery of upper limb function after stroke (Rapisarda et al. 1996; Escudero et al. 1998; Cruz Martinez et al. 1999; Shelton and Reding, 2001; Schiemanck et al. 2008; DeVetten et al. 2010; Radlinska et al. 2010; Puig et al. 2011). The algorithm...
assesses the effects of stroke on motor pathway integrity, beginning with a simple clinical evaluation of upper limb impairment and advancing to neurophysiological and neuroimaging measures if required. The aim of this study was to test and refine the PREP algorithm by comparing the patient stratification predicted by the algorithm with the stratification produced by an unbiased cluster analysis of upper limb function.

### Subjects and methods

#### Participants

Participants with ischaemic stroke were prospectively enrolled within 3 days of stroke onset from a single centre. Patients were eligible for the study if they were aged 18 years or older, had hemiparesis or hemiplegia, and had no prior history of stroke. Patients were excluded if they had bilateral or posterior circulation infarcts, homonymous hemianopia and cognitive or communication impairment or pre-existing conditions precluding informed consent or compliance with study assessments. Other exclusion criteria included pregnancy, a history of seizures and contraindications to MRI. The study was approved by the regional ethics committee, and all participants provided written informed consent in accordance with the Declaration of Helsinki.

#### Procedures

The SAFE score was determined 72 h after stroke onset, by grading shoulder abduction and finger extension of the affected upper limb using the Medical Research Council grades and summing these two scores (range 0–10). The Fugl-Meyer scale (Fugl-Meyer et al., 1975) and the National Institutes of Health stroke scale (Kasner, 2006) were used to evaluate upper limb impairment and stroke severity, respectively, 2 weeks after stroke. Upper limb function was evaluated with the Action Research Arm Test (ARAT), which is a reliable and validated measure that uses timed tests of a range of activities related to daily living (Van der Lee et al., 2001). ARAT scores were obtained at 2, 6, 12 and 26 weeks. The primary end-point for the study was the ARAT score at 12 weeks.

Participants received a standardized dose of upper limb rehabilitation beginning 2 weeks after stroke, consisting of ~30 min of physiotherapy and/or occupational therapy delivered every week day for 4 weeks. The content and intensity of each therapy session was determined by the treating therapists, who were blinded to all algorithm measures and clinical assessments for each participant. Participants also received standard care throughout the 26-week study. Clinical assessments were performed by therapists not involved in patient rehabilitation and who were blinded to the results of the SAFE score, TMS and diffusion-weighted imaging assessments.

TMS was used to test the functional integrity of the ipsilesional corticomotor pathway 2 weeks after symptom onset (mean 12.6, standard deviation (SD) 4.2 days). MEPs were recorded from the extensor carpi radialis muscle of the affected upper limb, using standard surface sEMG techniques. Recording electrodes (Ambu) were placed over the extensor carpi radialis in a belly-tendon montage and the reference electrode (Red Dot, 3 M) was placed over the lateral epicondyle of the humerus. Signals were sampled at 2 kHz, amplified (1000 gain), filtered (20–1000 Hz) and stored for offline analysis using Signal software (CED). Magnetic stimuli were delivered using a 70-mm figure-of-eight coil connected to a MagStim 200 unit. The coil was oriented to induce posterior-to-anterior current flow in the ipsilesional M1. The patient was categorized as MEP-positive (MEP+) if MEPs with a peak-to-peak amplitude ≥50 μV could be elicited on at least four out of eight consecutive trials while the muscle remained at rest. If this criterion was not met with stimuli delivered at maximum intensity, the patient was categorized as MEP-negative (MEP−).

MRI was used to evaluate the structural integrity of the posterior limb of the internal capsules 2 weeks after symptom onset (mean 10.8, SD 4.5 days). T1-weighted and diffusion-weighted images were acquired with a Siemens 1.5 T Avanto scanner. Axial T1-weighted images had 1.0 × 1.0 × 1.0 mm voxels, a 256-mm field of view, repetition time = 11 ms and echo time = 4.94 ms. Diffusion-weighted images had 1.8 × 1.8 × 3.0 mm voxels, a 230-mm field of view, b = 2000 s/mm², repetition = 6700 ms, echo time = 101 ms, 30 gradient directions and two averages. The mean fractional anisotropy (FA) was calculated within each posterior limb of the internal capsule by warping a template posterior limb of the internal capsule volume of interest to the patient’s images. The structural integrity of the posterior limb of the internal capsules was quantified by calculating an asymmetry index from the mean fractional anisotropy values: fractional anisotropy asymmetry index = (FAcontra − FAipsi)/(FAcontra + FAipsi) (Stinear et al., 2007). All image

### Table 1 Recovery definitions and examples of feasible rehabilitation goals

<table>
<thead>
<tr>
<th>Recovery</th>
<th>Definition</th>
<th>Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>The patient has the potential to return to normal or near-normal hand and arm function within 12 weeks.</td>
<td>Rehabilitation could focus on task-specific therapy in order to facilitate a return to full or near-full use of the hand and arm in activities of daily living.</td>
</tr>
<tr>
<td>Notable</td>
<td>The patient has the potential to be using their affected hand and arm in most activities of daily living within 12 weeks, though normal function is unlikely.</td>
<td>Rehabilitation could focus on strength, coordination and fine motor control, in order to maximize recovery of function and minimize compensation with the other hand.</td>
</tr>
<tr>
<td>Limited</td>
<td>The patient has the potential to have some movement in their affected hand and arm within 12 weeks, but it is unlikely to be used functionally for activities of daily living.</td>
<td>Rehabilitation could focus on reducing impairment by strengthening the paretic upper limb and improving active range of motion, in order to promote adaptation and incorporation of the affected upper limb in daily activities wherever possible.</td>
</tr>
<tr>
<td>None</td>
<td>The patient can expect to have minimal movement in their affected hand and arm, with little improvement at 12 weeks.</td>
<td>Rehabilitation could focus on prevention of secondary complications, such as spasticity and shoulder instability, and reducing disability by learning to complete activities of daily living with the unaffected hand and arm.</td>
</tr>
</tbody>
</table>

Participants received a standardized dose of upper limb rehabilitation beginning 2 weeks after stroke, consisting of ~30 min of physiotherapy and/or occupational therapy delivered every week day for 4 weeks. The content and intensity of each therapy session was determined by the treating therapists, who were blinded to all algorithm measures and clinical assessments for each participant. Participants also received standard care throughout the 26-week study. Clinical assessments were performed by therapists not involved in patient rehabilitation and who were blinded to the results of the SAFE score, TMS and diffusion-weighted imaging assessments.

TMS was used to test the functional integrity of the ipsilesional corticomotor pathway 2 weeks after symptom onset (mean 12.6, standard deviation (SD) 4.2 days). MEPs were recorded from the extensor carpi radialis muscle of the affected upper limb, using standard surface sEMG techniques. Recording electrodes (Ambu) were placed over the extensor carpi radialis in a belly-tendon montage and the reference electrode (Red Dot, 3 M) was placed over the lateral epicondyle of the humerus. Signals were sampled at 2 kHz, amplified (1000 gain), filtered (20–1000 Hz) and stored for offline analysis using Signal software (CED). Magnetic stimuli were delivered using a 70-mm figure-of-eight coil connected to a MagStim 200 unit. The coil was oriented to induce posterior-to-anterior current flow in the ipsilesional M1. The patient was categorized as MEP-positive (MEP+) if MEPs with a peak-to-peak amplitude ≥50 μV could be elicited on at least four out of eight consecutive trials while the muscle remained at rest. If this criterion was not met with stimuli delivered at maximum intensity, the patient was categorized as MEP-negative (MEP−).

MRI was used to evaluate the structural integrity of the posterior limb of the internal capsules 2 weeks after symptom onset (mean 10.8, SD 4.5 days). T1-weighted and diffusion-weighted images were acquired with a Siemens 1.5 T Avanto scanner. Axial T1-weighted images had 1.0 × 1.0 × 1.0 mm voxels, a 256-mm field of view, repetition time = 11 ms and echo time = 4.94 ms. Diffusion-weighted images had 1.8 × 1.8 × 3.0 mm voxels, a 230-mm field of view, b = 2000 s/mm², repetition = 6700 ms, echo time = 101 ms, 30 gradient directions and two averages. The mean fractional anisotropy (FA) was calculated within each posterior limb of the internal capsule by warping a template posterior limb of the internal capsule volume of interest to the patient’s images.
Statistical analysis

Recovery of upper limb function was confirmed by analysing ARAT scores with repeated measures ANOVA and four levels of the factor time (2, 6, 12, 26 weeks). Next, the algorithm was used to predict each participant’s level of upper limb recovery at 12 weeks. A univariate ANOVA with four levels of the factor recovery (complete, notable, limited, none) was used to see whether the algorithm’s predictions classified patients into four non-overlapping levels of upper limb function at 12 weeks. The analysis was repeated with a refined boundary for fractional anisotropy asymmetry index.

Cluster analyses were undertaken to provide a hypothesis-free classification of patients according to ARAT at 12 weeks, with no consideration of baseline measures. We used k-means cluster analyses to group the ARAT scores at 12 weeks into two, three, four and five clusters. We then inspected the results to determine the number of clusters that provided optimal resolution. Assuming that the minimal clinically important difference in ARAT score is 12 points (Lang et al., 2008), we adopted a criterion that adjacent cluster centres should be separated by at least 12 points in order to be clinically meaningful. Univariate ANOVA indicated descriptively the cluster analyses’ goodness-of-fit. The participant allocations produced by the four-cluster model were used as the basis for comparison with the algorithm’s allocations, allowing calculation of the positive and negative predictive power, and specificity and sensitivity of the algorithm.

The potential for using clinical and diffusion-weighted imaging measures alone to predict ARAT score at 12 weeks was also investigated. Linear regression was used to explore correlations between ARAT at 12 weeks and ARAT, Fugl-Meyer and fractional anisotropy asymmetry index at 2 weeks and SAFE score at 72 h. Correlational analysis was not undertaken for MEP data as these are binarized.

An alpha of 0.05 was adopted for statistical significance. Means and SDs are provided in the text except when 95% confidence intervals (CIs) are indicated.

Results

Forty participants (16 males; median age 70, range 31–91 years) with ischaemic stroke [13 (33%)] dominant hemisphere] were enrolled in the study between April 2010 and September 2011 (Table 2 and Fig. 2). As expected, ARAT score increased over time for the group \[F(3,36) = 46.18, P < 0.0001; \text{Fig. 3A}\]. The potential upper limb recovery at 12 weeks was predicted for each participant according to the proposed PREP algorithm (Table 3). The univariate ANOVA indicated a main effect of recovery on ARAT score at 12 weeks \[F(3,36) = 44.29, P < 0.0001\]. However, the 95% CI for the ‘limited’ and ‘none’ recovery stratifications overlapped, indicating that the initial fractional anisotropy asymmetry index boundary of 0.25 was sub-optimal. We then evaluated a refined version of the algorithm, with a more conservative boundary of fractional anisotropy asymmetry index = 0.15. This produced non-overlapping 95% CIs between all four stratifications and again the ANOVA was highly significant \[F(3,36) = 51.42, P < 0.0001\]. The partial \(\eta^2\) values indicate that the refined PREP algorithm explained more of the variance than the proposed PREP algorithm. The recovery trajectory of all participants, grouped by their stratification with the refined algorithm, is shown in Fig. 3B.

Cluster analysis

Analyses that grouped participants into two to five clusters produced descriptive ANOVA that were highly significant (all \(P < 0.0001\)). The two-cluster analysis corresponded to binarized MEP categories at baseline (MEP+, MEP−). The three-cluster
was 73%.

The positive predictive power of the algorithm was 88% (14 of the 16 positive predictions by the algorithm were true) and the negative predictive power was 83% (20 of the 24 negative predictions by the algorithm were true). For complete recovery, the specificity of the algorithm was 88% and the sensitivity was 73%.

Table 3 Comparison of the proposed and refined PREP algorithms

<table>
<thead>
<tr>
<th></th>
<th>Proposed PREP algorithm</th>
<th>Refined PREP algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
<td>F(3,36) = 44.29</td>
<td>F(3,36) = 51.42</td>
</tr>
<tr>
<td>P-value</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Partial $\eta^2$</td>
<td>0.787</td>
<td>0.811</td>
</tr>
<tr>
<td>ARAT at 12 weeks mean (95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete</td>
<td>54.9 (50.2–57.0)</td>
<td>54.9 (50.4–57.0)</td>
</tr>
<tr>
<td>Notable</td>
<td>43.1 (38.4–47.8)</td>
<td>43.1 (38.6–47.5)</td>
</tr>
<tr>
<td>Limited</td>
<td>14.4 (6.0–22.9)</td>
<td>21.3 (11.1–31.6)</td>
</tr>
<tr>
<td>None</td>
<td>1.0 (0–11.9)</td>
<td>2.2 (0–10.2)</td>
</tr>
</tbody>
</table>

One-way ANOVA was used to determine the effect of predicted upper limb recovery (complete, notable, limited, none) on ARAT score at 12 weeks. There was a significant effect when the proposed PREP algorithm was used to predict upper limb recovery. However, the CIs for the ‘limited’ and ‘none’ categories overlapped, indicating suboptimal classification by the algorithm. We then refined the algorithm by lowering the fractional anisotropy asymmetry index boundary between ‘limited’ and ‘none’, from 0.25 to 0.15. Subsequent ANOVA showed that the refined algorithm explained more of the variance (higher partial $\eta^2$) and produced non-overlapping CIs for the ‘limited’ and ‘none’ categories.

There was excellent correspondence between the four-cluster analysis of ARAT score at 12 weeks and predictions made at baseline using the PREP algorithm (Fig. 4B). Recovery was better than predicted for four patients and worse than predicted for three patients (Patients A, B and C; Fig. 4B). Patients A and B were able to complete all of the ARAT tasks, but their movements were slowed by their advanced age (90 years) and shoulder pain, respectively, limiting their scores on the timed ARAT. Patient C was the only participant with a pure premotor cortex lesion.

Correlations

The SAFE score at 72 h, ARAT score at 2 weeks and Fugl-Meyer score at 2 weeks all positively correlated with ARAT score at 12 weeks (SAFE: $r = 0.79$, $P < 0.001$; ARAT: $r = 0.85$, $P < 0.001$; FM2: $r = 0.86$, $P < 0.001$; Fig. 5A–C). Similarly, fractional anisotropy asymmetry index at 2 weeks was negatively correlated with ARAT score at 12 weeks (fractional anisotropy asymmetry index: $r = -0.61$, $P < 0.001$; Fig. 5D). The SAFE score and fractional anisotropy asymmetry index measures at 2 weeks were of little utility in predicting outcomes for individual patients on their own (as indicated by the 95% CI of the regressions) and could only predict ARAT score at 12 weeks when combined in a stepwise manner according to the PREP algorithm (Fig. 4).

Discussion

This study provides strong preliminary support for the use of the refined PREP algorithm to predict potential upper limb recovery in individual patients with stroke. A SAFE score at 72 h can identify those patients with the potential for ‘complete’ recovery of upper limb function. Patients with a SAFE score <8 at 72 h require further investigation at a later time point in order to accurately predict their recovery of upper limb function. TMS can then be
Figure 4. Cluster analyses of ARAT scores at 12 weeks identified groups inherent in the data, for comparison with the PREP algorithm predictions. (A) Symbols and vertical lines represent cluster centres and ranges. Grouping ARAT scores into four clusters produced the best resolution while maintaining at least 12 points between cluster centres (calibration bar MCID = minimal clinically important difference). The two-cluster analysis corresponded to binarized MEP categories at baseline (MEP+ , MEP−). The three-cluster analysis did not uniquely classify patients along any potentially useful tripartite combinations of baseline measures. The five-cluster analysis produced inadequate separation between adjacent cluster centres (<12 points). (B) The stratification of each patient by the PREP algorithm was checked against the group membership of each patient produced by the four-cluster analysis. Each symbol represents a patient belonging to one of four inherent clusters in the 12-week ARAT score. Each symbol is coloured according to their potential for upper limb recovery predicted by PREP. There is excellent correspondence between the groupings produced by the unbiased cluster analysis and the predictions made at 2 weeks using PREP. The algorithm predicted Patients A and B had the potential for ‘complete’ recovery, however they achieved a ‘notable’ recovery due to advanced age (Patient A) and shoulder pain (Patient B). Patient C’s predicted potential for ‘notable’ recovery wasn’t realized, probably due to extensive premotor cortex damage.
used to identify patients with ‘notable’ potential for recovery, and MRI is needed to distinguish between patients with ‘limited’ or ‘none’ recovery potential. The PREP algorithm may support tailored selection of upper limb rehabilitation goals and therapies (Table 1) and be used to stratify participants in clinical studies, based on the residual connectivity of key motor pathways of individual patients.

The PREP algorithm performed well, with 39/40 participants recovering at least as well as expected. We confirmed that participants with a SAFE score \( \geq 8 \) within 72 h of stroke make a ‘complete’ recovery of upper limb function within 12 weeks. While two participants in this category had lower-than-expected ARAT scores due to slow upper limb movement, their recovery can be considered clinically complete as they were able to successfully perform all of the tested activities. Participants with a SAFE score <8 who were MEP+ had a ‘notable’ recovery of upper limb function. For participants without MEPs, we confirmed that diffusion-weighted imaging is required to distinguish between those whose recovery is ‘limited’ or ‘none’, in keeping with recent reports that TMS has higher positive predictive value (Jang et al., 2010; Kwon et al., 2011).

All participants received a standardized dose of upper limb rehabilitation. It is possible that patients with ‘complete’ predicted recovery can realize their potential with a lower dose of upper limb therapy, but this remains to be directly tested. The PREP algorithm may be particularly useful for identifying patients whose potential for ‘notable’ upper limb recovery may otherwise go unrecognized and unrealized. Participants with a predicted recovery level of ‘none’ had more severe initial upper limb impairment, but clinical scores could not distinguish them from similarly impaired patients with ‘limited’ or ‘notable’ predicted recovery (Fig. 3B). Patients with a predicted recovery level of ‘none’ may not derive functional benefits from daily upper limb therapy but will still require rehabilitation and education in order to prevent secondary complications such as spasticity and shoulder pain.

The PREP algorithm is designed for efficiency and economy, by starting with simple bedside measures and only using more advanced techniques to resolve uncertainty for subsets of patients. In the current study, the algorithm required TMS assessment for

---

**Figure 5** Correlation analyses. The relationships between individual baseline measures and upper limb function 12 weeks after stroke are strong, but variable from patient to patient. Regression lines are shown with 95% CI. ARAT score at 12 weeks as a function of (A) the SAFE score at 72 h (\( r = 0.79; P < 0.001 \)), (B) ARAT score at 2 weeks (\( \text{ARAT}_2; r = 0.85; P < 0.001 \)), (C) Fugl–Meyer score at 2 weeks (\( \text{FM}_2; r = 0.86; P < 0.001 \)) and (D) fractional anisotropy asymmetry index at 2 weeks (\( r = -0.61; P < 0.001 \)).
60% of patients and MRI assessment for 20% of patients. The TMS assessment is relatively simple and inexpensive. If MEPs are present, there is no need to proceed to the more costly MRI scan. Note that the fractional anisotropy asymmetry index should not be calculated from MRI scans acquired within the first 5 days after stroke, as the measurement of ipsilesional fractional anisotropy is confounded by tissue oedema at this early stage (Sotak, 2002). However, the TMS assessment can be completed in the first 5 days, and the result will render a subsequent MRI scan unnecessary in a large proportion of patients.

This study refines the PREP algorithm by finding that the boundary between ‘limited’ and ‘none’ recovery stratifications is defined by a fractional anisotropy asymmetry index value of 0.15, rather than 0.25 as originally proposed for patients at the chronic stage of recovery (Stinear et al., 2007). Ipsilesional fractional anisotropy values are higher at the subacute stage and decrease with progressive degeneration of white matter integrity over the following months (Sotak, 2002). Therefore, for a given stroke the fractional anisotropy asymmetry index ‘point of no return’ is likely to have a lower value at the subacute than the chronic stage.

The group data indicated strong positive relationships between clinical measures made at 72 h (SAFE score) and at 2 weeks (ARAT and Fugl-Meyer scores) and the ARAT score at 12 weeks (Fig. 5) as expected. Similarly, there was a strong negative relationship between the structural integrity of the posterior limb of the internal capsule (fractional anisotropy asymmetry index) and ARAT score at 12 weeks. However, there was still a wide range of potential ARAT scores at 12 weeks for any given baseline measure of an individual patient. For example, patients with a SAFE score of 0 subsequently had ARAT scores across the full range, from 0 to 57 (Fig. 5A). Baseline clinical or diffusion-weighted imaging measures in isolation could not therefore be used to accurately predict the level of upper limb recovery at 12 weeks. Accurate prediction of individual outcomes was only possible when these measures were combined sequentially.

One of the potential limitations of this study is its sample size, with a small subset of patients being categorized as having ‘limited’ or ‘none’ recovery potential. Further work is needed to continue refining the algorithm’s boundary between these two categories. The external validity of this study may also be limited as the inclusion and exclusion criteria may reduce applicability to the broader stroke population. We have shown that PREP can be used in the absence of severe cognitive or communication impairments, or other significant co-morbidities. These factors may prevent the patient from realizing their full potential for motor recovery and need to be considered alongside PREP predictions when planning rehabilitation for each patient. The predicted potential for upper limb recovery was not attained in one patient (Patient C, Fig. 4B), whose recovery was ‘none’ despite being MEP+. This patient had an isolated and complete premotor cortex lesion and while the ipsilesional corticospinal tract was intact, he was unable to voluntarily move the distal affected upper limb. This could be due to a generalized loss of input to M1 from premotor cortical areas known to be important for movement preparation (Graziano, 2006). This indicates that cortical lesions ‘upstream’ of M1 may also limit the potential for recovery of motor function.

In conclusion, the PREP algorithm appears to predict the potential for upper limb recovery after stroke in individual patients. If confirmed in future studies, the algorithm may be used to manage patient and therapist expectations and set realistic upper limb rehabilitation goals based on the individual patient’s capacity for motor recovery. The PREP algorithm could also be used to stratify patients by identifying those who are more or less likely to respond in clinical trials of new upper limb therapies. Further work is needed to test the algorithm in patients with haemorrhagic or previous stroke and to explore the potential clinical and economic benefits of using PREP in clinical practice.

Acknowledgements

The authors thank Patricia Bennett, Marie-Claire Smith, Alison Elston and Anne Ronaldson for assistance with participant recruitment and data collection.

Funding

Health Research Council of New Zealand; Julius Brendel Trust and Stroke Foundation of New Zealand (Northern Region).

References


