

# Breaking Proportional Recovery After Stroke

Merav R. Senesh, MSc<sup>1</sup>  and David J. Reinkensmeyer, PhD<sup>1</sup>

Neurorehabilitation and  
Neural Repair  
2019, Vol. 33(11) 888–901  
© The Author(s) 2019  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/1545968319868718  
journals.sagepub.com/home/nnr



## Abstract

People with hemiparesis after stroke appear to recover 70% to 80% of the difference between their baseline and the maximum upper extremity Fugl-Meyer (UEFM) score, a phenomenon called proportional recovery (PR). Two recent commentaries explained that PR should be expected because of mathematical coupling between the baseline and change score. Here we ask, If mathematical coupling encourages PR, why do a fraction of stroke patients (the “nonfitters”) not exhibit PR? At the neuroanatomical level of analysis, this question was answered by Byblow et al—nonfitters lack corticospinal tract (CST) integrity at baseline—but here we address the mathematical and behavioral causes. We first derive a new interpretation of the slope of PR: It is the average probability of scoring across remaining scale items at follow-up. PR therefore breaks when enough test items are discretely more difficult for a patient at follow-up, flattening the slope of recovery. For the UEFM, we show that nonfitters are most unlikely to recover the ability to score on the test items related to wrist/hand dexterity, shoulder flexion without bending the elbow, and finger-to-nose movement, supporting the finding that nonfitters lack CST integrity. However, we also show that a subset of nonfitters respond better to robotic movement training in the chronic phase of stroke. These persons are just able to move the arm out of the flexion synergy and pick up small blocks, both markers of CST integrity. Nonfitters therefore raise interesting questions about CST function and the basis for response to intensive movement training.

## Keywords

stroke, rehabilitation, projections and predictions, movement, rating scales, proportional recovery

## Introduction

The proportional recovery (PR) rule was first observed in the context of upper extremity (UE) movement using the Fugl-Meyer (FM) score,<sup>1</sup> a widely used impairment-based measure in stroke rehabilitation research, and then confirmed in larger samples.<sup>2–4</sup> Samples of patients were evaluated at 1 week and at follow-up (typically 3 or 6 months) and most (the so-called “fitters”) recovered on average 70% to 80% (depending on the study) of the difference between their baseline and UEFM score and the maximum UEFM score (Figure 1A). The rule is:

$$\overline{\Delta S} = m(T - S_o) \quad (1)$$

where  $\overline{\Delta S}$  is the mean change in score for all subjects with a baseline score of  $S_o$ ,  $T$  is the maximum score (66 for the UEFM), and  $m \approx 0.7$  to  $0.8$  is the slope. Other subjects, termed “nonfitters,” achieved substantially less recovery than that predicted from the PR rule (Figure 1A). The PR rule has subsequently been shown for measures of lower limb impairment, visuospatial neglect, aphasia, and success at reaching pellets in a rodent stroke model (reviewed in

Hawe et al<sup>5</sup>), leading to the possibility that it reflects a fundamental biological mechanism of recovery.

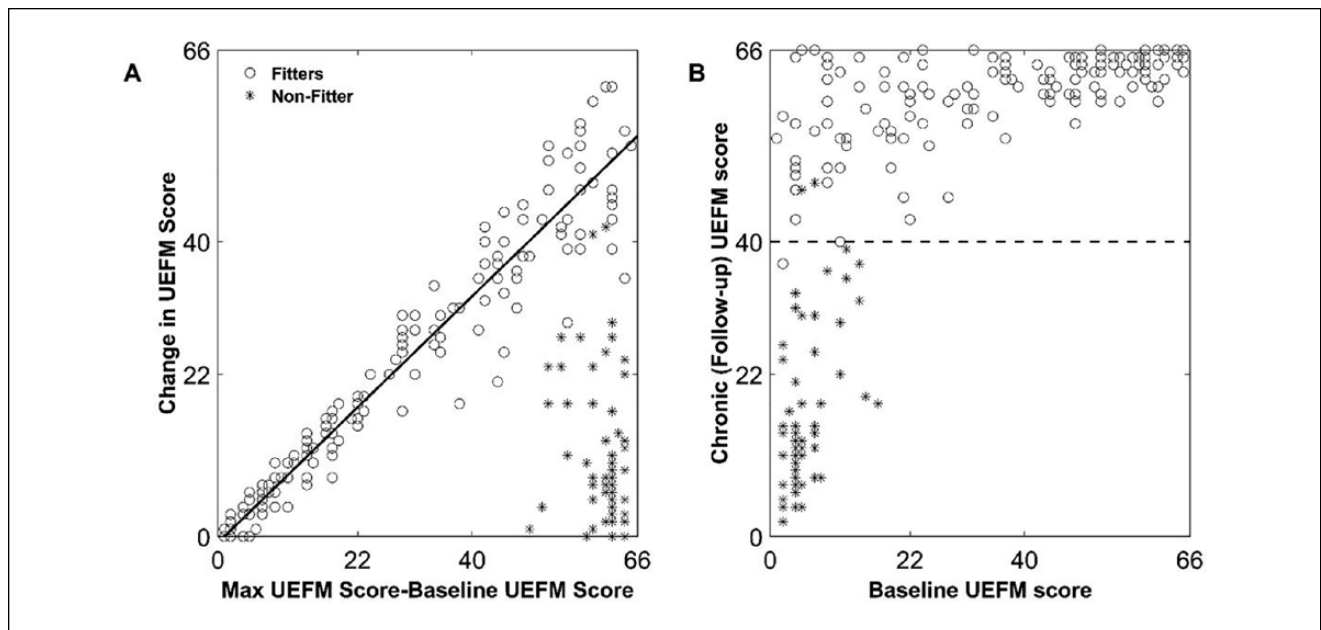
Predictions of stroke recovery are multifactorial and highly variable,<sup>6</sup> so the finding of such a regularity, especially one driven by a single impairment-based input and present for a wide variety of outcome measures, is striking and even mathematically suspicious, or in the words of Hawe et al,<sup>5</sup> “too good to be true.” Indeed, Hawe et al<sup>5</sup> and Hope et al<sup>7</sup> recently used statistical analysis and simulation to demonstrate that a PR-like relationship arises even if outcomes

<sup>1</sup>University of California at Irvine, Irvine, CA, USA

Supplementary material for this article is available on the *Neurorehabilitation & Neural Repair* website along with the online version of the article.

## Corresponding Author:

David J. Reinkensmeyer, PhD, Department of Mechanical and Aerospace Engineering, Department of Anatomy and Neurobiology, Department of Biomedical Engineering, Department of Physical Medicine and Rehabilitation, University of California at Irvine, 4200 Engineering Gateway, Irvine, CA 92697-3975, USA.  
Email: dreinken@uci.edu



**Figure 1.** Proportional recovery of the upper extremity Fugl-Meyer (UEFM) score after stroke. These plots combine data that we digitized from reported data of 214 subjects enrolled in 2 studies.<sup>2,3</sup> Baseline scores were measured within 2 weeks poststroke and follow-up at 6 months (chronic stage) (A) For “fitters” (o), the average change in UEFM score is proportional to how many points a patient has left to recover (ie, the maximum UEFM score of 66 minus the baseline UEFM score). “Nonfitters” (\*) show decreased recovery. (B) By plotting the same data another way (ie, showing the relationship between chronic [follow-up] UEFM score and baseline UEFM score), it is apparent that 99% of patients with a chronic UEFM score less than 40 (dashed line) in the chronic stage are nonfitters.

are unpredictable and recovery is random. Mathematical coupling between  $\Delta S$  and  $S_o$  (since  $\Delta S = m(T - S_o)$ ), high variability of  $S_o$  compared with  $S_f$ , and ceiling effects all contribute to PR. These phenomena are known in statistical theory and difficult to correct.<sup>5</sup>

But if mathematical coupling encourages PR, then why does PR *not* arise for a particular subset of stroke patients—the “nonfitters”—who recover less than that predicted by PR? At the neuroanatomical level of analysis, this question has already been elegantly answered: Lack of corticospinal tract (CST) integrity strongly predicts nonfitters.<sup>3,4,8-10</sup> The first to demonstrate this were Byblow et al<sup>3</sup> who used motor-evoked potentials (MEPs) and found that individuals who at baseline evaluation did not have an MEP response (ie, who were “MEP–”) were highly likely to be the nonfitters. Stinear et al<sup>4</sup> confirmed this predictive value of MEP in a larger sample.

Here, we sought to gain mathematical and behavioral insight into the rule breakers of PR by analyzing the rule in terms of the way the UEFM score is constructed—as a sum of subscores on a series of test items—and determining how the ability to score on those items at follow-up mathematically influences PR. We also queried the effect of rehabilitation for fitters and nonfitters by analyzing data from robotic UE training.

## Preliminary Observations

We digitized reported data from 214 subjects enrolled in 2 studies<sup>2,3</sup> for analysis here. The inclusion criteria for both studies were (see Table S1 in the Supplemental Material for more detail): (1) first ever ischemic stroke: anterior circulation<sup>2</sup>/monohemispheric,<sup>3</sup> (2) assessments within 2 weeks of stroke, (3) no severe cognitive or communication impairment, (4) no severe mobility and daily activity impairment (premorbid Barthel Index of  $\geq 19$ ),<sup>2</sup> (5) no homonymous hemianopia or blindness, no visuospatial neglect or complete somatosensory loss,<sup>3</sup> and (6) no evidence for cerebellar stroke. Only 7% out of 4082 consecutive stroke inpatients met the above criteria in the study by Winters et al<sup>2</sup> (36% were unwilling to participate, died, or transferred to other centers). However, in a latter study, Stinear et al<sup>4</sup> conformed the same result for a wider stroke population which included people with recurrent stroke and intracerebral hemorrhage.

The second preliminary observation has to do with identifying nonfitters. It is not possible to reliably identify nonfitters using their baseline UEFM alone.<sup>2,9-11</sup> Individuals with lower initial UEFM are not necessarily the nonfitters, as can be seen in Figure 1A. However, importantly for our analysis here, it is possible to identify nonfitters when they reach the

*chronic* stage of recovery using an UEFM threshold of about 40. Ninety-nine percent of patients with a chronic UEFM score less than 40 in the chronic stage are nonfitters, as seen in Figure 1B. That is, fitters nearly always recover to an UEFM greater than 40. We will use this observation below in several analyses.

## Interpreting the PR Rule in Terms of Scale Construction

We first analyze PR for scales that are calculated by summing subscores on a set of items (such as the UEFM score). We show that PR arises robustly for such scales (as expected because of mathematical coupling<sup>5,7</sup>), and that the slope has a probabilistic interpretation related to scale construction. We also show that the local slope is not constant for previously reported UEFM PR data, consistent with this probabilistic interpretation. Finally, PR is broken (producing nonfitters) when a substantial number of scale items are discretely more difficult to score on at follow-up for some patients. Finally, we identify the specific UEFM test items that nonfitters have the lowest probability of being able to perform.

## Constructing an Outcome Measure

Assume we evaluate recovery at two time points (baseline and follow-up) using a measure that is composed by adding  $T$  subscores on a set of items. Assume each item can be scored 0 (cannot perform) or 1 (performs well). Then,  $T$  is the maximum total score a patient can achieve on the measure. Assume a patient has achieved a score  $S_o$  at baseline by scoring a 1 on items 1 to  $S_o$ . Then, there are  $T - S_o$  items left for him/her to score 1 on at the follow-up, assuming patients do not regress on items (ie, they will score a 1 on any item at follow-up that they scored 1 on at baseline).

Now, assume that at the follow-up the subscore for each remaining item  $i$  is a random variable  $t_i$  that takes the value 1 with probability  $p_i$  and the value 0 with probability  $1 - p_i$ . We assume that the  $p_i$  are functions of  $(T - S_o)$  (the number of remaining items = the number of points the patient has left to achieve following the baseline assessment, if items are scored 0 and 1). For example, we expect that patients who are better recovered at baseline will have a higher probability of scoring well on the remaining items at the follow-up. We will call  $p_i(T - S_o)$  the “item recovery probability function” (IRPF) since  $p_i$  is the probability of scoring 1 on the remaining item  $i$  given a score of  $S_o$  at baseline. Then, the increase in score at follow-up is also a random variable:

$$\Delta S = t_1 + t_2 + \dots + t_{T-S_o} \quad (2)$$

Now, imagine designing an outcome measure—what should one design the IRPF’s look like? A common strategy in scale design is to observe a sample of patients at baseline and then follow them; this appears to be how the UEFM scale was developed.<sup>12</sup> Based on observations and some theoretical basis for a patient’s evolving capabilities (such as a progression out of abnormal synergies for the UEFM scale<sup>12</sup>), one then designs a set of test items such that there is a finite probability that some fraction of subjects will be able to do each item at follow-up. Ideally, these probabilities will vary across items, creating a range of items from easy to difficult. Indeed, in one formal approach to scale development, Rasch analysis, scales are designed so that the probabilities of achieving items are spaced across the range of ability measured over a sample of recovered subjects.<sup>13</sup> When Rasch analysis is applied to the UEFM scale, it shows that the UEFM items span a range of difficulties and can be rank-ordered in terms of those difficulties.<sup>14</sup>

Figure 2A shows a set of theoretical IRPFs with the number of test items  $T = 33$ , matching the UEFM scale. The lower curves represent more difficult items. The probabilities increase with increasing baseline score (ie, decreasing remaining items to score on), capturing the intuition that patients who are better recovered at baseline have a higher chance of scoring on items at follow-up.

## Simulating Recovery

Given a set of IRPFs, we can simulate recovery. To produce a total score equal to that possible with the UEFM scale (ie, 66 points), we assumed the 33 items could be scored a 0 or a 2. We used a uniform initial distribution of baseline scores  $S_o$  for 120 patients. Their predicted recovery exhibited a PR-like behavior (Figure 2A, right column).

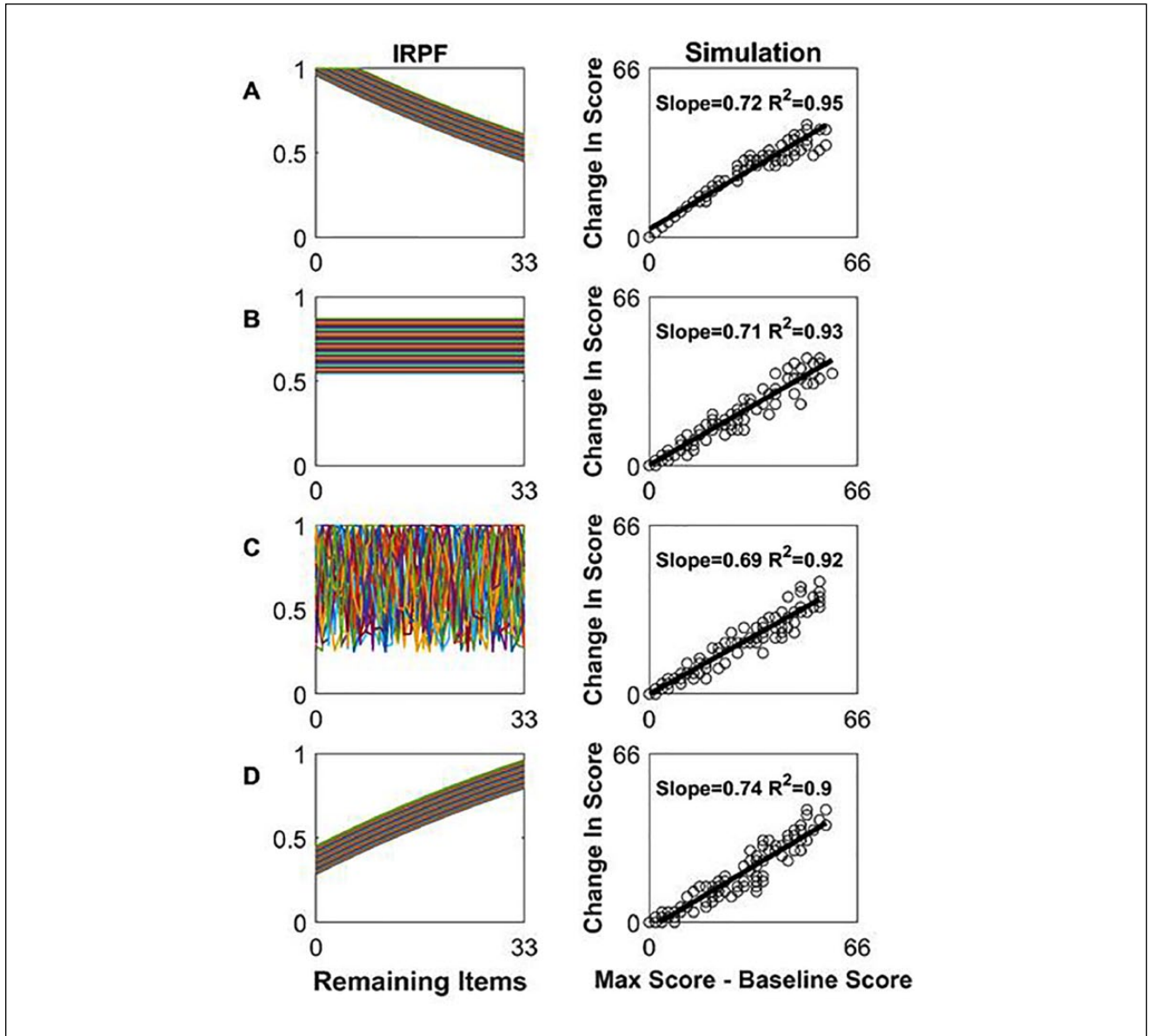
How dependent is the PR-like behavior on the shape of the IRPFs? We ran the simulation with several types of IRPFs:  $p_i$  = a random constant that varied with item number  $i$  but not  $S_o$  (Figure 2B);  $p_i$  = random number that

varied with  $i$  and  $S_o$  (Figure 2C); and  $p_i = \min \left( \frac{-x}{e^T} + b, 1 \right)$

(i.e. nonlinear decay in recovery probability) (Figure 2D). In each case, we obtained a PR-like rule (Figure 2 right column), as would be expected due to mathematical coupling.

## Bending PR: The Local Slope

We can gain insight into why PR rises robustly for a wide range of IRPFs by estimating the slope of the PR curve. Returning to the case when only scores of 0 and 1 are possible, the mean increase in score from Equation (2) is:



**Figure 2.** Proportional recovery (PR) arises robustly from simulations using a variety of item recovery probability functions (IRPFs). Left column: Examples of IRPFs, the functions that define the probability of scoring on an item at follow-up given the number of remaining items left to score on. Each line is the IRPF for a different item, for a test composed of 33 items scored 0 or 2. Right column: Recovery results from simulating 120 patients. As can be seen, a variety of IRPFs produce a PR-like relationship, which is expected given mathematical coupling.

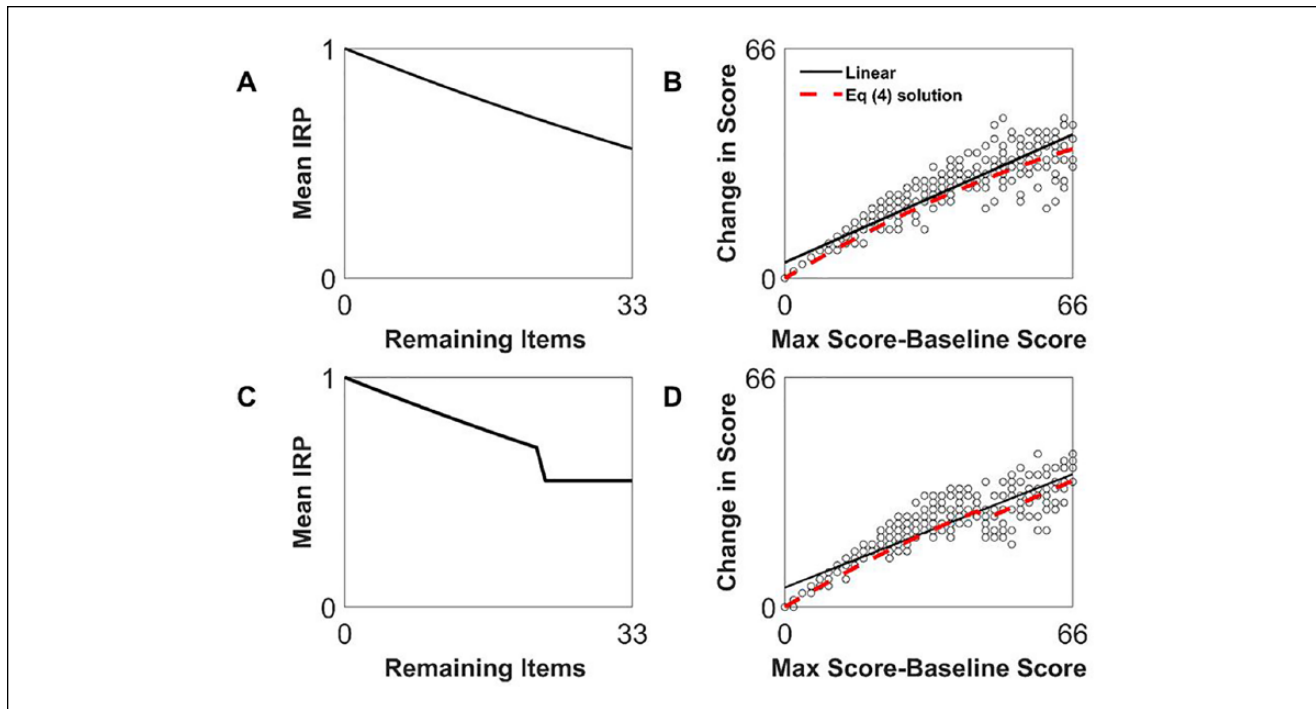
$$\begin{aligned}
 \overline{\Delta S} &= \bar{t}_1 + \bar{t}_2 + \cdots + \bar{t}_{T-S_o} \\
 &= p_1 + p_2 + \cdots + p_{T-S_o} \\
 &= \frac{(p_1 + p_2 + \cdots + p_{T-S_o})(T - S_o)}{T - S_o} \\
 &= \bar{p}(T - S_o)(T - S_o)
 \end{aligned} \quad (3)$$

$\bar{p}(T - S_o)$  is the average item recovery probability across the remaining items for which the subject has not yet scored

1. Equation 3 potentially restates the PR rule of Equation 1, but proportionality will hold exactly only if  $\bar{p}(T - S_o)$  is constant across  $S_o$ . If  $\bar{p}(T - S_o)$  is not constant, then the local slope of the relationship is

$$m = \bar{p}(T - S_o) + \frac{\partial \bar{p}}{\partial (T - S_o)}(T - S_o) \quad (4)$$

Equation 4 can be solved assuming a constant  $m$ , resulting in a family of functions shifted by an arbitrary constant. Thus, an infinite number of  $\bar{p}(T - S_o)$  curves produce an



**Figure 3.** Bending and breaking proportional recovery. Shown is the relationship of average item recovery probability function (IRPF, left column) and recovery (right column). In the top row, the average IRPF is smooth but not constant, and produces a nonlinear relationship that is nevertheless well fit by a linear regression (solid line,  $R^2 = 0.88$ ). The red dashed line shows the predicted fit from Equation 4. In the bottom row, the average IRPF is discontinuous, but the resulting recovery relationship does not look discontinuous, but rather just somewhat nonlinear, and again is still well fit by linear regression ( $R^2=0.83$ ), although the local slope varies as Equation 4 (red dashed line).

exact PR rule. Can we gain insight into scales that would *not* be expected to produce an exact PR?

First, note that if the IRPFs vary smoothly across  $S_o$ ,  $\bar{p}(T - S_o)$  will also vary smoothly. In this case, even if  $\bar{p}(T - S_o)$  is not constant, one still gets a PR-like relationship, with the local slope varying according to Equation 4 (Figure 3).

If  $\bar{p}(T - S_o)$  is (first-order) unsmooth, then  $\frac{\partial \bar{p}}{\partial (T - S_o)}$  jumps, causing the instantaneous slope  $m(S_o)$  to jump (Figure 3C). But slope discontinuities are masked by the stochastic nature of the IRPFs (Figure 3D).

It is well established in statistics that it is difficult to discern local slope variations with noisy data and/or a limited sample size. Furthermore, global linear regression, which is the only method that appears to have been applied to detect PR, is ill-suited for detecting nonlinearity.<sup>15</sup> We therefore suggest that the proportionality of PR, besides arising from mathematical coupling, is also an approximation that arises partly because of the smoothness of the average IRPF, and partly because global linear regression cannot detect local slope variations.

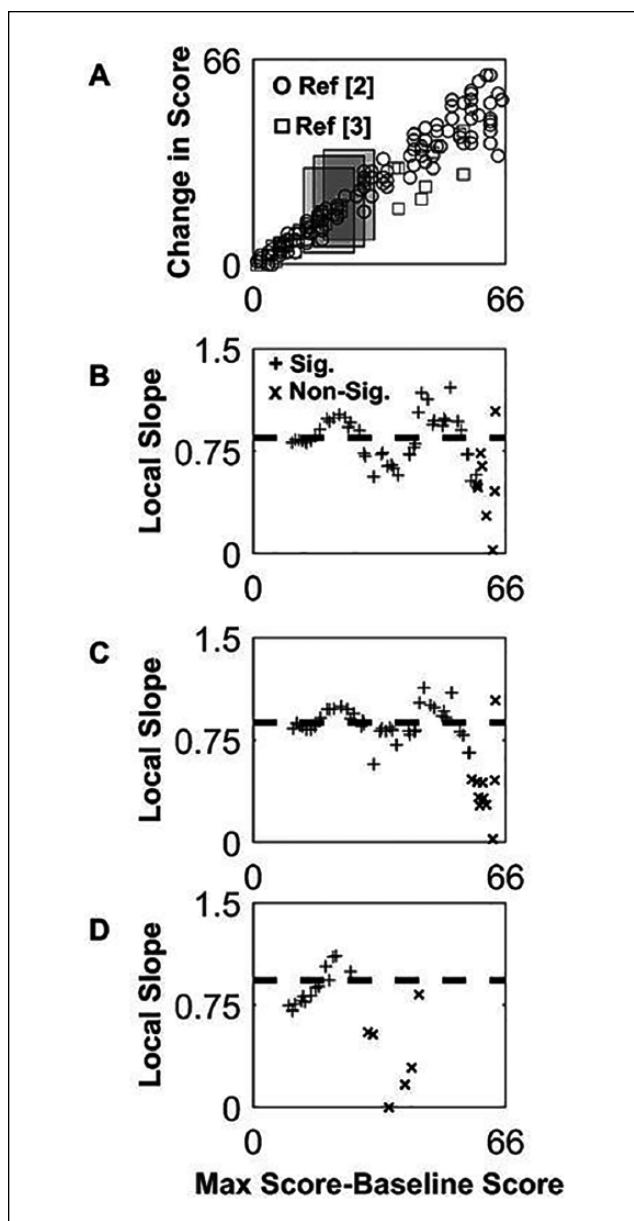
We used local linear regression to estimate if the local slope of PR varies for different impairment levels. The local slope is the slope of a line fitted to a subset of the data,

where the subset of data is selected as the data points (ie, patients) that fall in a window of size 20 along the  $x$ -axis. We shifted the window along the  $x$ -axis to estimate the slope for different data subsets.

Figure 4 shows that the local slope estimate varies from 0.0 to 1.0, in a similar pattern for 2 different data sets and is not constant as would be expected for exact proportionality across impairment levels (Equation 3). Note that the local slope estimated in the windows over impairment level of 46 to 66 (ie, initial UFEM < 20) is smaller (Figure 4C). This is consistent with the expectation that average recovery probability  $\bar{p}(T - S_o)$  will decrease with decreasing initial score  $S_o$ —that is, subjects with lower baseline scores have a lower probability of scoring on remaining items at follow-up (resulting in a lower local slope). Also, the global regression slope closely matches the local slope for the patients who are less impaired at baseline—these subjects have smaller recovery variance (partly due to a ceiling effect), which drives the global regression.

### Breaking PR: Nonfitters

What are the mathematical conditions needed to create non-fitters who break the PR rule, given the various statistical factors working positively toward creating an appearance of



**Figure 4.** Estimating the local slope of recovery using local linear regression. (A) Upper extremity Fugl-Meyer (UEFM) recovery data from Winters et al.<sup>2</sup> (circles) and Byblow et al.<sup>3</sup> (squares) shown with sample windows (20 UEFM points wide) used to create the local estimate of slope (gray rectangles). For the data points (ie, subjects) falling within each window, we performed a regression to estimate the slope for those data points alone (ie, the “local slope”). (B) Local slope estimated with the linear regression over each window for the data combined from.<sup>2,3</sup> (C) Local slope from Winters et al.<sup>2</sup> alone, with window size of 20 UEFM points. (D) Local slope from Byblow et al.<sup>3</sup> alone, with a window size of 20 UEFM points. Plus (+) signs show points for which the  $P$ -value of the local regression  $< .05$ , and “x” where the  $P > .05$ . Horizontal dashed line shows slope from global regression. Note that the global regression slope result is dictated by the data from the individuals with higher baseline scores because their data are less variable.

PR? Working with our simulator, we found that if we set recovery probabilities for a substantial number of items to a low value for a specific set of subjects with a baseline score below some score (ie, we made a set of items extremely difficult for those low-level subjects, Figure 5A and C), then those subjects recovered poorly in the simulation, and they appeared as nonfitters (Figure 5B and D).

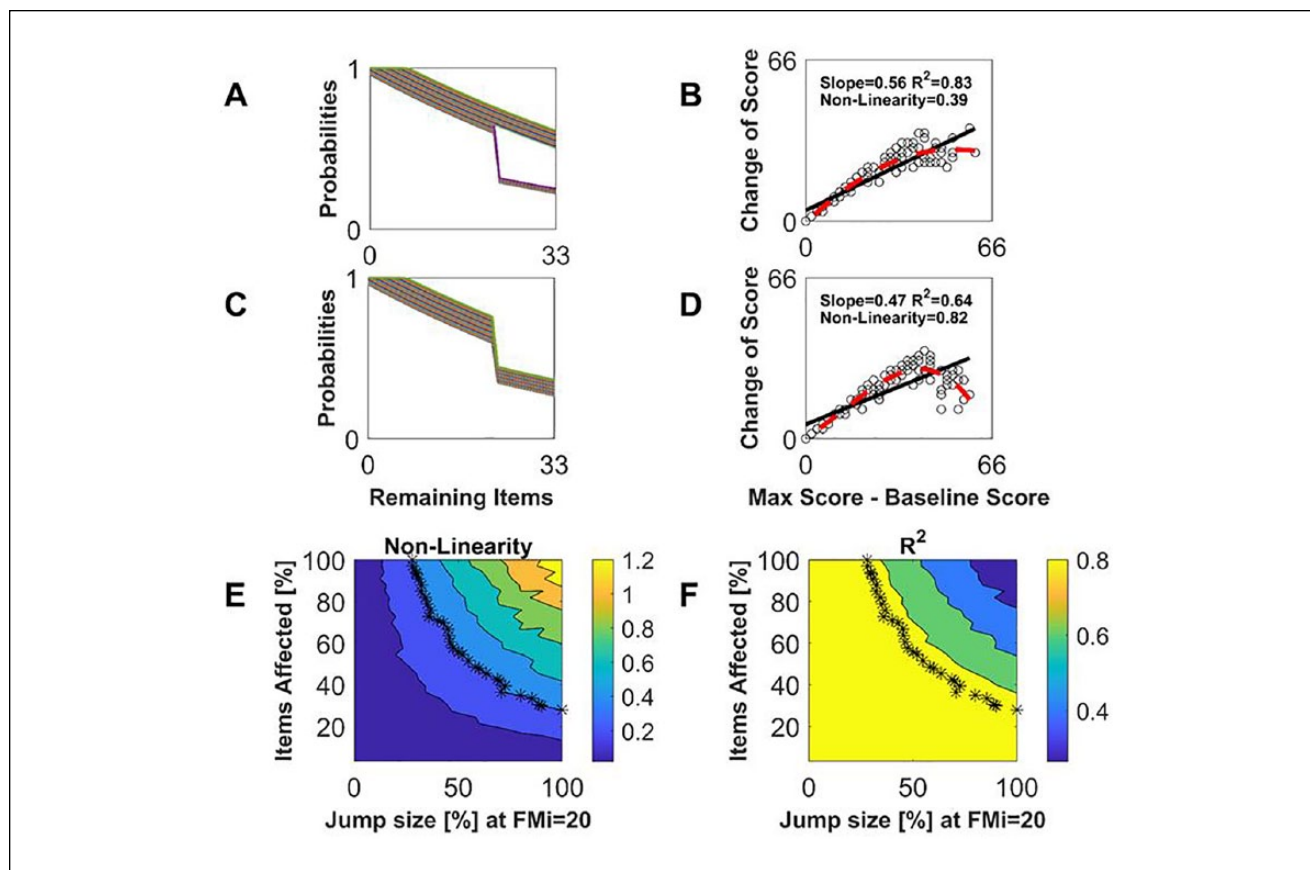
To quantify how substantial the IRPF discontinuities have to be to produce a visible nonlinearity, we simulated the situation where  $P\%$  of the IRPFs, starting with the most difficult items, exhibited a jump decrease by  $D\%$  at  $S_o = 20$ , varying  $P$  and  $D$  (Figure 5). Large values of  $D$  mean those items are nearly impossible for a subset of patients to score on at follow-up. We analyzed the  $R^2$  and nonlinearity of the resulting PR-like rule, where we defined nonlinearity  $NL$  as the root mean square deviation of a third-degree polynomial fit to the data away from a linear fit.<sup>16,17</sup> Nonlinearity became visually obvious when  $NL$  approached  $\sim 0.4$  (Figure 5B).

Achieving an  $NL = 0.4$  required at least  $\sim 50\%$  jumps in at least  $\sim 50\%$  of the items (Figure 5B and E).  $R^2$  values were always greater than 0.8 for simulations with  $NL = 0.4$  (Figure 5F), confirming that linear regression statistics were poor at signaling the emerging nonlinearity.<sup>15</sup>

### Very Difficult UEFM Items for Nonfitters

Given that very difficult items can “break” PR, what are the very difficult items for the nonfitters of the UEFM at follow-up? As observed above, 99% of patients with a UEFM score less than 40 in the chronic stage are nonfitters. We used UEFM measurements from 319 chronic stroke patients reported in several studies<sup>18–25</sup> to rank-order items from low to high probability of scoring on that item (either 1 or 2), using the estimated probabilities at UEFM = 40 to create the rank ordering. We found that individuals with chronic UEFM  $< 40$  (ie, nonfitters) had the lowest probabilities on the items that required distal coordinated movement (Figure 6). These “very difficult” items tested the ability to move distal joints (wrist circumduction, wrist stability, forearm pronation/supination, wrist flexion/extension, grasp A—that is, “pretend you are holding the handle of a brief case” and grasp B—that is, abducting the thumb to grasp a piece of paper), to actively flex the shoulder through the extreme of its range of motion (shoulder flexion from  $90^\circ$  to  $180^\circ$ ) without bending the elbow, and to make repeated high-speed targeted reaching movements without vision (ie, the finger-to-nose movement item of the UEFM test). The 2 remaining wrist/hand items were easy/medium, but the easy one only required mass flexion of the hand, and the medium one required flexion of the wrist, without requiring control over other joints or active extension of the fingers or wrist. That is, “in-synergy” UEFM items were relatively easy for





**Figure 5.** Relationship between item recovery probability function (IRPF) nonlinearities and nonfitters. A great enough jump nonlinearity in a great enough number of the IRPFs creates a nonlinearity in the proportional recovery (PR)–type relationship, even though the nonlinearities are smoothed by the summation associated with calculating a total score. (A) Sample IRPFs with 50% jump nonlinearities in 30% of the items. (B) PR plot resulting from IRPFs in A. The nonlinearity measure was quantified as

$$NL = \frac{\sqrt{\int_{X_L}^{X_R} (g(x) - f(x))^2 dx}}{X_L - X_R} \quad \text{where } g(x) \text{ is the best-fit third-order polynomial (dashed line) and } f(x) \text{ is the linear PR model (solid line).}$$

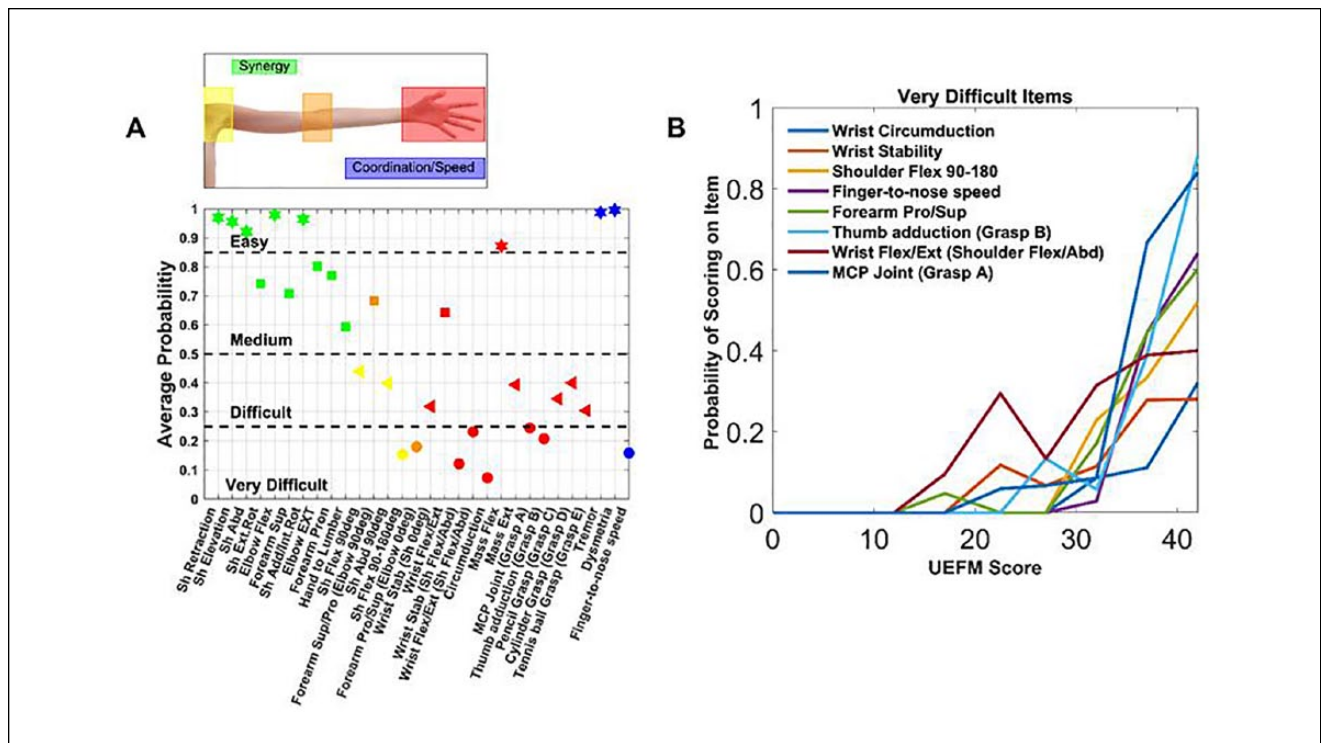
Nonlinearity started becoming visually apparent with  $NL \geq 0.4$ . (C) Sample IRPFs with 40% jump nonlinearities in all items. (D) PR plot resulting from IRPFs in C. (E) Results from simulations where  $P\%$  of the 33 IRPFs, starting with the most difficult items, exhibited a jump decrease by  $D\%$  at  $S_o = 20$ , varying  $P$  and  $D$ . (F)  $R^2$  values for the same simulations as in E. The asterisks in E and F show simulation runs where  $L = 0.4$ , the value where we judged nonlinearity to be visually apparent. In panel F, one can see that simulations with  $NL = 0.4$  always produced  $R^2$  values greater than 0.8—that is,  $R^2$  is a poor indicator of nonlinearity.

these nonfitters to achieve, in contrast to the items requiring distal coordinated movement (Figure 6).

## Summary

To summarize, we derived a new interpretation of the slope of PR: It is the average probability of scoring across remaining scale items at follow-up. The proportionality of PR, besides arising from mathematical coupling, is also an approximation that arises because of the smoothness of the average IRPF, and because global linear regression cannot detect local slope variations. The local slope estimate varies substantially in a consistent pattern for 2

different published UEFM recovery data sets—a result not predicted by mathematical coupling, but understandable from the scale construction perspective. Finally, what is required to overcome these mathematical phenomena that encourage the appearance of PR is that a patient must exhibit discretely lower probability of scoring on a substantial subset of test items at follow-up. We showed that UEFM nonfitters are most unlikely to score on test items related to wrist/hand dexterity, extreme active shoulder range of motion, and finger-to-nose speed at follow-up. This is consistent with the idea that nonfitters lack sufficient CST, since CST is thought to be the neural substrate for these types of hand and arm movements.<sup>26,27</sup>



**Figure 6.** Analysis of difficulty on scoring on items of the upper extremity Fugl-Meyer (UEFM) assessment. (A) Average probability of scoring either 1 or 2 on each UEFM item for nonfitters taken as a group (ie, subjects with follow-up UEFM scores <40). The 3 reflex test items were excluded as all recorded subjects scored 1 or 2 on these items. We categorized item difficulty based on the probability of scoring on the item. We also highlight the type of movement tested: synergistic movement (green), shoulder (yellow), elbow (orange), wrist + hand (red), and UE tremor/coordination/speed (blue). (B) Probability of scoring on the very difficult items as a function of chronic UEFM score. Items are listed from most difficult to easiest in the legend, top to bottom.

## How Do Nonfitters Respond to Rehabilitation?

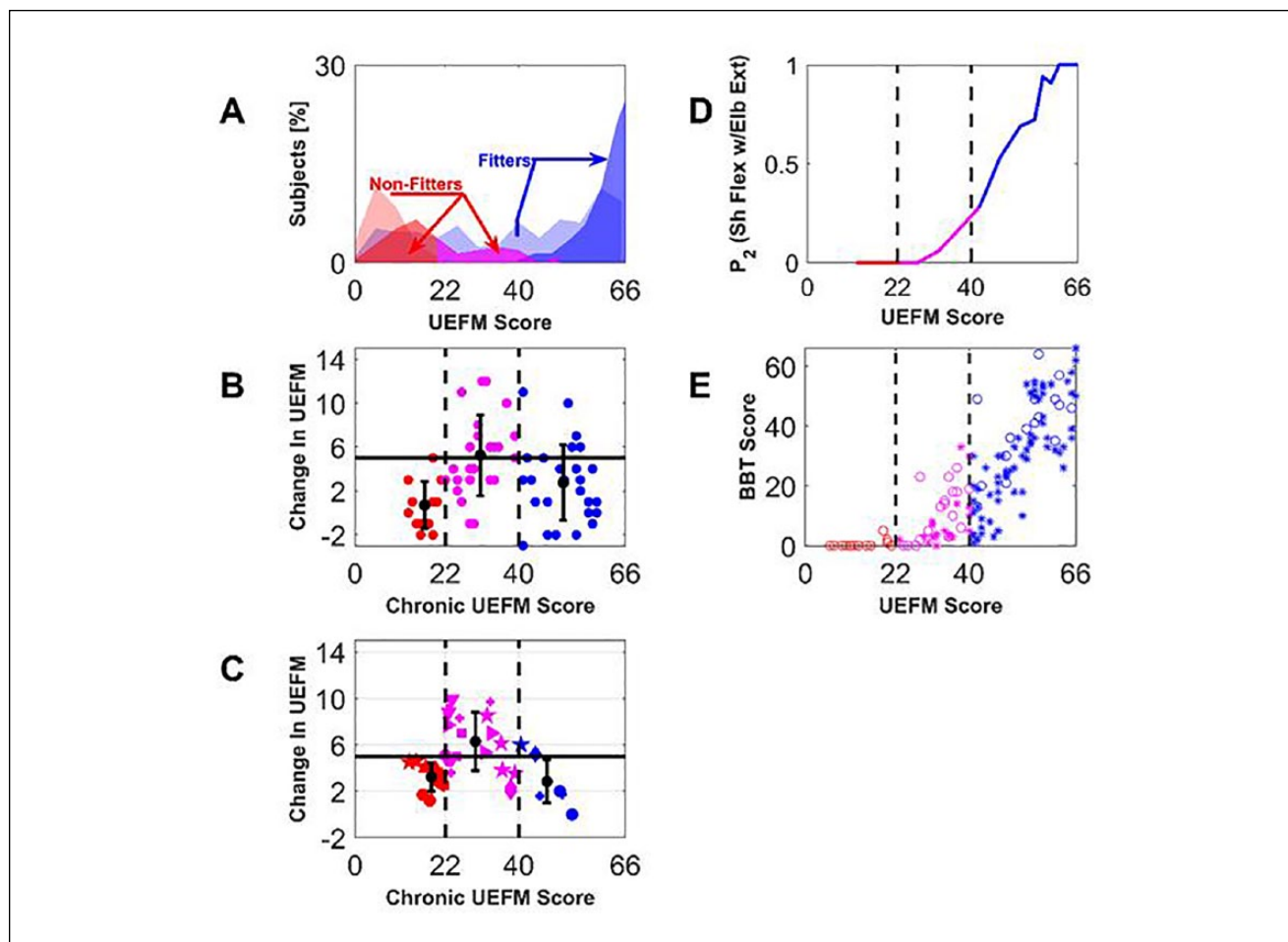
One possible implication of PR is that rehabilitation therapy does not matter because patients follow the PR prediction despite varying amounts of therapy.<sup>3,28</sup> However, Hawe et al<sup>5</sup> highlighted that there is substantial variability in UEFM recovery even for the fitters. The number of fitters displaying  $\approx 70\%$  recovery is near chance levels. Nonfitters have even greater variability of recovery (Figure 1), leading us to hypothesize that they are particularly responsive to the varying amounts of movement training they receive during rehabilitation.

To test this hypothesis, we returned to an analysis of UEFM response to robotic UE therapy in the chronic phase of stroke.<sup>29</sup> We grouped patients based on UEFM scores at study entry that fell into 3 zones: [0-21], [22-40], or [41-66]. Interestingly, these groups correspond to poorly recovered nonfitters, better-recovered nonfitters, and fitters, respectively, in the acute phase studies (Figure 7A, data taken from Winters et al<sup>2</sup> and Byblow et al<sup>3</sup>). Here we make the assumption that people who qualified for a robotic therapy study in the chronic phase and had a

UEFM score <40 would have been classified as nonfitters had they been tracked from the acute phase. It may be that there is a different pattern of upper extremity recovery that has not yet been identified for the type of patient who would not have qualified for PR studies, but then qualified for a chronic robotic therapy study (see Supplemental Material). However, in examining the inclusion/exclusion criterion for the chronic robotic therapy studies, we found they were similar to the PR studies (Table S1, Supplemental Material). Where the criteria deviated (such as not requiring a first-ever anterior circulation stroke), a recent study examining a broader, more clinically relevant sample of patients with stroke, including those with intracerebral hemorrhage and previous stroke found the same pattern of upper extremity PR (the pattern of fitters and nonfitters) and the same predictive value of MEP.<sup>4</sup> Thus, we feel the assumption that participants with chronic UEFM score <40 would have been classified as nonfitters is reasonable, although it should be verified in future research.

For each group, we calculated the mean reported change in UEFM score due to robotic rehabilitation. Nonfitters with mean baseline UEFM scores in the range





**Figure 7.** Analysis of a special group of nonfitters who respond better to intensive upper extremity (UE) movement training provided in the chronic phase of stroke. (A) Distribution of patients (% out of total 214 patients reported in Winters et al<sup>2</sup> and Byblow et al<sup>3</sup>) at baseline (light colors) and follow-up (dark colors), “fitters” (blue) and “nonfitters” (red, poor recovery; magenta, better recovered). (B) Change in upper extremity Fugl-Meyer (UEFM) score after UE robotic movement training as reported in our previous work.<sup>29</sup> The better-recovered nonfitters showed an increased UEFM response to robotic movement training (ANOVA,  $P < .001$ , with follow-up  $t$  tests, group 1-2 comparison  $P = .001$ , group 2-3 comparison  $P = .015$ ; group 1-3 comparison  $P = .03$ ). (C) Mean change in UEFM score reported for control and experimental groups in 18 robotic rehabilitation clinical trials with chronic stroke patients (studies<sup>21,31-47</sup> taken from the systematic review<sup>30</sup>). Treatment groups included conventional therapy, functional electrical stimulation, and robot-assisted arm training. Matching symbols indicate groups from the same studies. For B and C, the horizontal solid line shows the minimal clinically important difference (MCID) for the UEFM test, the vertical dashed lines denote how we subdivided the subjects into 3 UEFM zones. The groups with mean starting UEFM score in the middle zone again had a significantly higher mean change in UEFM score (ANOVA,  $P < .001$ , with follow-up  $t$  tests, group 1-2 comparison  $P < .001$ , group 2-3 comparison  $P = .0013$ ; group 1-3 comparison  $P = .56$ ). (D) Estimated probability to score “2” in the UEFM test item testing shoulder flexion from 0° to 90° with the elbow fully extended, plotted as a function of UEFM score. As described in the text, this is based on data from 319 subjects from several studies.<sup>18-25</sup> Subjects with UEFM in the range [22-40] (ie, the rehabilitation responsive nonfitters) are the ones just able to perform this out-of-flexion-synergy movement. (E) Box and Blocks Test (BBT) score as a function of the UEFM based on data from several studies.<sup>22,25,50</sup> Subjects with UEFM in the range [22-40] also are the ones predicted to just able to grasp and release blocks on the BBT. Here, we identified the underlying relationship between motor performance on the UEFM and the BBT in order to estimate BBT scores from UEFM scores.

[22-40] benefited significantly more from rehabilitation delivered in the chronic phase after stroke (Figure 7B, ANOVA,  $P < .001$ , with follow-up  $t$  tests, group 1-2 comparison  $P = .001$ , group 2-3 comparison  $P = .015$ ; group 1-3 comparison  $P = .03$ ).

We sought to test this finding using data from a larger number of studies. We therefore analyzed the rehabilitation response of nonfitters using results from a set of randomized controlled trials of robotic and conventional rehabilitation therapy reported in a systematic review.<sup>30</sup>

We considered the studies that enrolled persons in the chronic phase after stroke, leaving 18 trials.<sup>20,31-47</sup> We sorted results from these studies into groups with reported mean UEFM scores at study entry that fell into 3 zones: [0-21], [22-40], or [41-66].

These groups again correspond to poorly recovered nonfitters, better recovered nonfitters, and fitters, respectively (Figure 7A, data taken from Winters et al<sup>2</sup> and Byblow et al<sup>3</sup>).

For each group we calculated the mean reported change in UEFM score due to rehabilitation of any kind (ie, robot or control therapy).

Participants with mean baseline UEFM scores in the range [22-40] (ie, the better-recovered nonfitters) again benefited significantly more from intensive movement practice in the chronic phase after stroke (Figure 7C, ANOVA,  $P < .001$ , with follow-up  $t$  tests, group 1-2 comparison  $P < .001$ , group 2-3 comparison  $P = .0013$ ; group 1-3 comparison  $P = .56$ ).

Note that the distribution of patients becomes increasingly bimodal as recovery progresses, with the better-recovered fitters in the “bottom of the well” of this distribution (Figure 7A). Thus, the better-recovered nonfitters are somewhat scarce: they are ~12% of all stroke patients enrolled in the PR trial data we scanned, and 45% of the nonfitters.

As described above, the lowest probability UEFM items for nonfitters are the ones that require hand and arm dexterity, consistent with lack of CST integrity. But perhaps these better-recovered, rehabilitation responsive nonfitters have return of CST integrity. From the perspective of abnormal synergies, a litmus test of CST integrity is the ability to flex the shoulder to 90° while keeping the elbow fully extended<sup>48,49</sup>—that is, the test of “out-of-flexion-synergy” movement on the UEFM. We plotted the probability of scoring a “2” on the test as a function of UEFM score, again for 319 chronic stroke patients.<sup>18-25</sup> Figure 7D shows that this probability becomes nonzero in the UEFM range of better recovered nonfitters, consistent some return of CST function.

Return of CST function should also be associated with return of hand function. The UEFM coarsely grades hand impairment and does not test hand function, so we turned to the Box and Blocks Test (BBT). The BBT requires individuals to pick up small blocks and place them over a divider and is thus a functional measure designed to be sensitive to the return of dexterity. We used published data from 140 subjects<sup>22,25,50</sup> with both BBT and FM scores to define the underlying relationship between these 2 motor scales. Using this relationship, we found that the better recovered nonfitters would be expected to exhibit a small amount of hand dexterity on the BBT (Figure 7E). Thus, the putative rehabilitation-responsive nonfitters, who are known to be MEP– at baseline due to the work of Byblow et al,<sup>3</sup> appear to be the patients in the chronic stage who can

begin to move the arm out of synergy and pick up small blocks, both markers of some CST integrity.

## Discussion: Point of View and Directions for Research

Mathematical coupling, as well as other statistical factors identified here, encourage the appearance of PR, and yet a fraction of stroke patients, the nonfitters, do not exhibit PR. We showed that the local slope of PR relates to test item difficulties. Thus, PR breaks when a patient has a discretely lower probability of scoring on a substantial fraction of test items at follow-up. For the UEFM, the lowest probability items are the ones that test wrist/hand dexterity, extreme active shoulder flexion, and high-speed finger-to-nose movement, consistent with the finding that nonfitters lack CST integrity.<sup>3,4,8-10</sup>

Winters et al<sup>2</sup> found that lack of finger extension within 72 hours poststroke was the best behavioral predictor of nonfitters. The present results highlight that nonfitters continue to do poorly on tests of hand dexterity, and also arm coordination, and this is why they overcome the influence of mathematical coupling.

However, better-recovered nonfitters appear to exhibit significantly enhanced UEFM response to robotic UE movement training, given the assumption that people who qualified for a robotic therapy study in the chronic phase and had an UEFM score  $<40$  would have been classified as nonfitters had they been tracked from the acute phase (see Supplemental Material for an analysis of this assumption). A related result was recently noted in an analysis of reaching rehabilitation in a rat stroke model in the subacute phase ( $n = 123$  rats).<sup>51</sup> A subset of rats classified as nonfitters responded better to reaching practice. Fitters recovered the ability to recover food pellets regardless of whether they performed reaching practice. Thus, intensive movement practice mattered for the nonfitters but not the fitters.

In the context of the UEFM, what could explain the heightened responsiveness of better-recovered nonfitters to movement practice? One possibility is that the UEFM scale is nonlinear, making it easier to score larger gains when starting from the middle range of the UEFM. Our analysis of the UEFM scale does not support this—it is more difficult to score on items when one starts from the middle range (see Figure 4). Furthermore, the UEFM has a very small floor effect and the ceiling effect is also small<sup>52,53</sup> and unlikely to constrain recovery for people with scores in the range 40 to 50, since the UEFM can go up to 66 and is linearly related to BBT in the range 40 to 50 (and beyond see Figure 7E). Yet subjects in this range of 40 to 50 nonetheless showed relatively decreased rehabilitation response (see Figure 7B and C). Another possibility is that individuals in the chronic phase of stroke experience a decay in their UEFM over time due to disuse of the UE. Thus, when

people in the chronic phase of stroke were enrolled in the movement training studies analyzed here, they may have simply overcome that decay. Studies that tracked motor function after stroke for years have indeed sometimes found a decay in UEFM or other UE motor scores,<sup>54,55</sup> but at other times have not.<sup>56,57</sup> Neither of these alternate explanations account for the related finding from the rodent study, since the outcome measure there was not the UEFM (it was pellets retrieved) and the rats received rehabilitation soon after the induced stroke, not in the chronic stage.

Do a subset of nonfitters therefore defy the expectation of a poorer response to UE rehabilitative movement training that has been associated with higher amounts of CST damage.<sup>58-60</sup> Perhaps these nonfitters rely on cortical plasticity to make use of residual pathways<sup>10</sup> such as brainstem pathways.<sup>48,61,62</sup> Another possibility is that these subjects have some residual CST integrity. Transcranial magnetic stimulation-based assays of CST integrity might miss detecting small amounts of residual CST early after stroke, or, alternatively, some level of neural reorganization might need time to occur and allow a small amount of residual CST to become functional. Our findings regarding the emergence of out-of-flexion-synergy arm movement and hand manipulation in the rehabilitation-responsive, nonfitter population support this idea, as these are behavioral markers of CST integrity. Byblow et al<sup>3</sup> found that 6 of the 11 initially MEP- participants recovered MEPs at later evaluations. Birchenall et al<sup>63</sup> also found recovery of MEP. In the study by Byblow et al,<sup>3</sup> MEP- participants also showed some recovery if fractional anisotropy of the CST was small, suggesting residual CST without initial MEP. This finding is supported by a later study of Buch et al<sup>11</sup> but not that of Doughty et al.<sup>64</sup> Stinear et al<sup>4</sup> also showed that MEP- patients had greater UEFM recovery if they had lower lesion load to the CST, consistent with the results reported by Feng et al.<sup>9</sup>

We recently found that chronic stroke patients did not respond to robotic hand movement training if they lacked somatosensory integrity.<sup>65</sup> Somatosensory integrity was indeed the strongest predictor of rehabilitation response in this context, stronger than CST integrity.<sup>66</sup> This highlights the fact that damage to other tracts and systems play a key role in determining response to intensive hand movement training. Somatosensory integrity might be required to effectively carry the feedback necessary to drive CST-related plasticity. In this scenario, perhaps it is the initially MEP- but bettered-recovered and rehabilitation-responsive nonfitters who have this somatosensory integrity.

If a small amount of CST, initially unresponsive to transcranial magnetic stimulation, serves as the basis for an increased response to movement training, it is unclear why people with presumably a greater amount CST integrity—the lower level fitters (with chronic UEFM in the range

40-50)—have a relatively reduced response to intensive movement training. Perhaps fitters make progress through spontaneous limb use, making any applied training appear less effective. In contrast, the special population of nonfitters might need to learn to make use of their small amounts of residual CST by performing especially large amounts of dedicated, practice-based exploration.<sup>67,68</sup> Work from our laboratory on robotic exoskeleton training is consistent with this idea, in that it found that people in the chronic phase of stroke who were MEP+, but with smaller rather than larger MEP, responded better to UE robotic movement training.<sup>69</sup> In any case, an important direction for future research is to determine how best to identify this putative special class of nonfitters as early as possible after stroke, since they may be the best candidates for intensive movement training, robotic or otherwise, if one is interested in selecting candidates who will have the largest UEFM change from the training.


### Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: David Reinkensmeyer has received payment for consulting and holds equity in Hocoma and Flint Rehabilitation Technologies, manufacturers of rehabilitation technology. The terms of these interests have been reviewed by the U.C. Irvine Conflict of Interest committee.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research was supported in part by grant NIH-2R01HD062744 from the National Center for Medical Rehabilitation Research at NIH.

### ORCID iD

Merav R. Senesh  <https://orcid.org/0000-0001-9273-1414>

### References

1. Prabhakaran S, Zarahn E, Riley C, et al. Inter-individual variability in the capacity for motor recovery after ischemic stroke. *Neurorehabil Neural Repair*. 2008;22:64-71. doi:10.1177/1545968307305302
2. Winters C, van Wegen EEH, Daffertshofer A, Kwakkel G. Generalizability of the proportional recovery model for the upper extremity after an ischemic stroke. *Neurorehabil Neural Repair*. 2015;29:614-622. doi:10.1177/1545968314562115
3. Byblow WD, Stinear CM, Barber PA, Petoe MA, Ackerley SJ. Proportional recovery after stroke depends on corticomotor integrity. *Ann Neurol*. 2015;78:848-859. doi:10.1002/ana.24472
4. Stinear CM, Byblow WD, Ackerley SJ, Smith M, Borges VM, Barber PA. Proportional motor recovery after stroke: implications for trial design. *Stroke*. 2017;48:795-798. doi:10.1161/STROKEAHA.116.016020

5. Hawe RL, Scott SH, Dukelow SP. Taking proportional out of stroke recovery. *Stroke*. 2019;50:204-211. doi:10.1161/STROKEAHA.118.023006
6. Boyd LA, Hayward KS, Ward NS, et al. Biomarkers of stroke recovery: consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. *Int J Stroke*. 2017;12:480-493. doi:10.1177/1747493017714176
7. Hope TMH, Friston K, Price CJ, Leff AP, Rotshtein P, Bowman H. Recovery after stroke: not so proportional after all? *Brain*. 2019;142:15-22. doi:10.1371/journal.pone.0147697
8. Koch P, Schulz R, Hummel FC. Structural connectivity analyses in motor recovery research after stroke. *Ann Clin Transl Neurol*. 2016;19:233-244. doi:10.1002/acn3.278
9. Feng W, Wang J, Chhatbar PY, et al. Corticospinal tract lesion load: an imaging biomarker for stroke motor outcomes. *Ann Neurol*. 2015;78:860-870. doi:10.1002/ana.24510
10. Zarahn E, Alon L, Ryan SL, et al. Prediction of motor recovery using initial impairment and fMRI 48 h poststroke. *Cereb Cortex*. 2018;12:2712-2721. doi:10.1093/cercor/bhr047
11. Buch ER, Rizk S, Nicolo P, Cohen LG, Schnider A, Guggisberg AG. Predicting motor improvement after stroke with clinical assessment and diffusion tensor imaging. *Neurology*. 2016;86:1924-1925. doi:10.1212/WNL.0000000000002675
12. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. *Scand J Rehabil Med*. 1975;7:13-31.
13. McDowell I. *Measuring Health: A Guide to Rating Scales and Questionnaires*. 3rd ed. New York, NY: Oxford University Press; 2006. doi:10.1179/108331900786166731
14. Woodbury ML, Velozo CA, Richards LG, Duncan PW, Studenski S, Lai SM. Dimensionality and construct validity of the Fugl-Meyer assessment of the upper extremity. *Arch Phys Med Rehabil*. 2007;88:715-723. doi:10.1016/j.apmr.2007.02.036
15. Anscombe FJ. Graphs in statistical analysis. *Am Stat*. 1973;27:17-21. doi:10.1080/00031305.1973.10478966
16. Emancipator K, Kroll MH. A quantitative measure of nonlinearity. *Clin Chem*. 1993;39:766-772. doi:10.1002/prca.201090041
17. Kroll MH, Emancipator K. A theoretical evaluation of linearity. *Clin Chem*. 1993;39:405-413.
18. Dodakian L, McKenzie AL, Le V, et al. A home-based telerehabilitation program for patients with stroke. *Neurorehabil Neural Repair*. 2017;31:923-933. doi:10.1177/1545968317733818
19. Taheri H, Rowe JB, Gardner D, Chan V, Reinkensmeyer DJ, Wolbrecht ET. Robot-assisted guitar hero for finger rehabilitation after stroke. *Conf Proc IEEE Eng Med Biol Soc*. 2012;2012:3911-3917.
20. Housman SJ, Scott KM, Reinkensmeyer DJ. A randomized controlled trial of gravity-supported, computer-enhanced arm exercise for individuals with severe hemiparesis. *Neurorehabil Neural Repair*. 2009;23:505-514.
21. Milot MH, Spencer SJ, Chan V, et al. A crossover pilot study evaluating the functional outcomes of two different types of robotic movement training in chronic stroke survivors using the arm exoskeleton BONES. *J Neuroeng Rehabil*. 2013;10:112. doi:10.1186/1743-0003-10-112
22. Reinkensmeyer DJ, Wolbrecht ET, Chan V, Chou C, Cramer SC, Bobrow JE. Comparison of 3D, assist-as-needed robotic arm/hand movement training provided with Pneu-WREX to conventional table top therapy following chronic stroke. *Am J Phys Med Rehabil*. 2012;91:S232-S241.
23. Secoli R, Milot MH, Rosati G, Reinkensmeyer DJ. Effect of visual distraction and auditory feedback on patient effort during robot-assisted movement training after stroke. *J Neuroeng Rehabil*. 2011;8:21. doi:10.1186/1743-0003-8-21
24. Taheri H, Rowe JB, Gardner D, et al. Design and preliminary evaluation of the FINGER rehabilitation robot: controlling challenge and quantifying finger individuation during musical computer game play. *J Neuroeng Rehabil*. 2014;11:10. doi:10.1186/1743-0003-11-10
25. Wolbrecht ET, Rowe JB, Chan V, Ingemanson ML, Cramer SC, Reinkensmeyer DJ. Finger strength, individuation, and their interaction: relationship to hand function and corticospinal tract injury after stroke. *Clin Neurophysiol*. 2018;129:797-808. doi:10.1016/j.clinph.2018.01.057
26. Lemon RN. Descending pathways in motor control. *Annu Rev Neurosci*. 2008;31:195-218. doi:10.1146/annurev.neuro.31.060407.125547
27. Devanne H, Cohen LG, Kouchtir-Devanne N, Capaday C. Integrated motor cortical control of task-related muscles during pointing in humans. *J Neurophysiol*. 2002;87:3006-3017. doi:10.1152/jn.2002.87.6.3006
28. Krakauer JW, Carmichael ST, Corbett D, Wittenberg GF. Getting neurorehabilitation right: what can be learned from animal models? *Neurorehabil Neural Repair*. 2012;26:923-931. doi:10.1177/1545968312440745
29. See J, Dodakian L, Chou C, et al. A standardized approach to the Fugl-Meyer assessment and its implications for clinical trials. *Neurorehabil Neural Repair*. 2013;27:732-741. doi:10.1177/1545968313491000
30. Mehrholz J, Pohl M, Platz T, Kugler J, Elsner B. Electromechanical and robot-assisted arm training for improving activities of daily living, arm function, and arm muscle strength after stroke. *Cochrane Database Syst Rev*. 2018;9:CD006876. doi:10.1002/14651858.CD006876.pub5
31. Ang KK, Guan C, Phua KS, et al. Brain-computer interface-based robotic end effector system for wrist and hand rehabilitation: results of a three-armed randomized controlled trial for chronic stroke. *Front Neuroeng*. 2014;7:30. doi:10.3389/fneng.2014.00030
32. Conroy SS, Whittall J, Dipietro L, et al. Effect of gravity on robot-assisted motor training after chronic stroke: a randomized trial. *Arch Phys Med Rehabil*. 2011;92:1754-1761. doi:10.1016/j.apmr.2011.06.016
33. Daly JJ, Hogan N, Perepezko EM, et al. Response to upper-limb robotics and functional neuromuscular. *J Rehabil Res Dev*. 2005;42:723-736. doi:10.1682/JRRD.2005.02.0048
34. Hsieh Y, Wu C, Liao W, Lin K, Wu K, Lee C. Effects of treatment intensity in upper limb robot-assisted therapy for chronic stroke: a pilot randomized controlled trial. *Neurorehabil Neural Repair*. 2011;25:503-511. doi:10.1177/1545968310394871

35. Klamroth-Marganska V, Blanco J, Campen K, et al. Three-dimensional, task-specific robot therapy of the arm after stroke: a multicentre, parallel-group randomised trial. *Lancet Neurol*. 2014;13:159-166. doi:10.1016/S1474-4422(13)70305-3
36. van der Lee JH, Wagenaar RC, Lankhorst GJ, Vogelaar TW, Devillé WL, Bouter LM. Forced use of the upper extremity in chronic stroke patients. *Stroke*. 1999;30:2369-2375. doi:10.1161/01.STR.30.11.2369
37. Hsieh YW, Lin KC, Horng YS, Wu CY, Wu TC, Ku FL. Sequential combination of robot-assisted therapy and constraint-induced therapy in stroke rehabilitation: a randomized controlled trial. *J Neurol*. 2014;261:1037-1045. doi:10.1007/s00415-014-7345-4
38. Hwang CH, Seong JW, Son DS. Individual finger synchronized robot-assisted hand rehabilitation in subacute to chronic stroke: a prospective randomized clinical trial of efficacy. *Clin Rehabil*. 2012;26:696-704. doi:10.1177/0269215511431473
39. Lo AC, Guarino PD, Richards LG, et al. Robot-assisted therapy for long-term upper-limb impairment after stroke. *N Engl J Med*. 2010;362:1772-1783. doi:10.1056/NEJMoa0911341
40. Burgar CG, Lum PS, Shor PC, Van der Loos HFM. Development of robots for rehabilitation therapy: the Palo Alto VA/Stanford experience. *J Rehabil Res Dev*. 2000;37:663-673.
41. McCabe J, Monkiewicz M, Holcomb J, Pundik S, Daly JJ. Comparison of robotics, functional electrical stimulation, and motor learning methods for treatment of persistent upper extremity dysfunction after stroke: a randomized controlled trial. *Arch Phys Med Rehabil*. 2015;96:981-990. doi:10.1016/j.apmr.2014.10.022
42. Susanto EA, Tong RK, Ockenfeld C, Ho NS. Efficacy of robot-assisted fingers training in chronic stroke survivors: a pilot randomized-controlled trial. *J Neuroeng Rehabil*. 2015;12:42. doi:10.1186/s12984-015-0033-5
43. Timmermans AA, Lemmens RJ, Monfrance M, et al. Effects of task-oriented robot training on arm function, activity, and quality of life in chronic stroke patients: a randomized controlled trial. *J Neuroeng Rehabil*. 2014;11:45. doi:10.1186/1743-0003-11-45
44. Volpe BT, Lynch D, Rykman-Berland A, et al. Intensive sensorimotor arm training mediated by therapist or robot improves hemiparesis in patients with chronic stroke. *Neurorehabil Neural Repair*. 2008;22:305-310. doi:10.1177/1545968307311102
45. Wu CY, Yang CL, Chuang LL, et al. Effect of therapist-based versus robot-assisted bilateral arm training on motor control, functional performance, and quality of life after chronic stroke: a clinical trial. *Phys Ther*. 2012;92:1006-1016.
46. Brokaw EB, Nichols D, Holley RJ, Lum PS. Robotic therapy provides a stimulus for upper limb motor recovery after stroke that is complementary to and distinct from conventional therapy. *Neurorehabil Neural Repair*. 2014;28:367-376. doi:10.1177/1545968313510974
47. Bustamante Valles K, Montes S, Madrigal Mde J, Burciaga A, Martínez ME, Johnson MJ. Technology-assisted stroke rehabilitation in Mexico : a pilot randomized trial comparing traditional therapy to circuit training in a robot/technology-assisted therapy gym. *J Neuroeng Rehabil*. 2016;13:83. doi:10.1186/s12984-016-0190-1
48. McPherson JG, Chen A, Ellis MD, Yao J, Heckman CJ, Dewald JPA. Progressive recruitment of contralesional cortico-reticulospinal pathways drives motor impairment post stroke. *J Physiol*. 2018;596:1211-1225. doi:10.1113/JP274968
49. McMorland AJC, Runnalls KD, Byblow WD. A neuroanatomical framework for upper limb synergies after stroke. *Front Hum Neurosci*. 2015;9:82. doi:10.3389/fnhum.2015.00082
50. Platz T, Pinkowski C, van Wijck F, Kim I-H, di Bella P, Johnson G. Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: a multi-centre study. *Clin Rehabil*. 2005;19:404-411. doi:10.1191/0269215505cr832oa
51. Jeffers MS, Karthikeyan S, Gomez-Smith M, et al. Does stroke rehabilitation really matter? Part B: an algorithm for prescribing an effective intensity of rehabilitation. *Neurorehabil Neural Repair*. 2018;32:73-83. doi:10.1177/1545968317753074
52. Lin JH, Hsu MJ, Sheu CF, et al. Psychometric comparisons of 4 measures for assessing upper-extremity function in people with stroke. *Phys Ther*. 2009;89:840-850. doi:10.2522/ptj.20080285
53. Hsueh IP, Hsu MJ, Sheu CF, Lee S, Hsieh CL, Lin JH. Psychometric comparisons of 2 versions of the Fugl-Meyer Motor Scale and 2 versions of the Stroke Rehabilitation Assessment of Movement. *Neurorehabil Neural Repair*. 2008;22:737-744. doi:10.1177/1545968308315999
54. Löfgren B, Nyberg L, Mattsson M, Gustafson Y. Three years after in-patient stroke rehabilitation: a follow-up study. *Cerebrovasc Dis*. 1999;9:163-170. doi:10.1159/000015948
55. Meyer S, Verheyden G, Brinkmann N, et al. Functional and motor outcome 5 years after stroke is equivalent to outcome at 2 months: follow-up of the collaborative evaluation of rehabilitation in stroke across Europe. *Stroke*. 2015;46:1613-1619. doi:10.1161/STROKEAHA.115.009421
56. Broeks JG, Lankhorst GJ, Rumping K, Prevo AJ. The long-term outcome of arm function after stroke: results of a follow-up study. *Disabil Rehabil*. 1999;21:357-364.
57. Feys H, De Weerd W, Verbeke G, et al. Early and repetitive stimulation of the arm can substantially improve the long-term outcome after stroke: a 5-year follow-up study of a randomized trial. *Stroke*. 2004;35:924-929. doi:10.1161/01.STR.0000121645.44752.f7
58. Stinear CM, Barber PA, Smale PR, Coxon JP, Fleming MK, Byblow WD. Functional potential in chronic stroke patients depends on corticospinal tract integrity. *Brain*. 2007;130(pt 1):170-180. doi:10.1093/brain/awl333
59. Riley JD, Le V, Der-Yeghiaian L, et al. Anatomy of stroke injury predicts gains from therapy. *Stroke*. 2011;42:421-426. doi:10.1161/STROKEAHA.110.599340
60. Cassidy JM, Tran G, Quinlan EB, Cramer SC. Neuroimaging identifies patients most likely to respond to a restorative stroke therapy. *Stroke*. 2018;49:433-438. doi:10.1161/STROKEAHA.117.018844

61. Baker SN. The primate reticulospinal tract, hand function and functional recovery. *J Physiol.* 2011;589(pt 23):5603-5612. doi:10.1113/jphysiol.2011.215160
62. Riddle CN, Edgley SA, Baker SN. Direct and indirect connections with upper limb motoneurons from the primate reticulospinal tract. *J Neurosci.* 2009;29:4993-4999. doi:10.1523/JNEUROSCI.3720-08.2009
63. Birchenall J, Térémetz M, Roca P, et al. Individual recovery profiles of manual dexterity, and relation to corticospinal lesion load and excitability after stroke—a longitudinal pilot study. *Neurophysiol Clin.* 2019;49:149-164. doi:10.1016/J.NEUCLI.2018.10.065
64. Doughty C, Wang J, Feng W, Hackney D, Pani E, Schlaug G. Detection and predictive value of fractional anisotropy changes of the corticospinal tract in the acute phase of a stroke. *Stroke.* 2016;47:1520-1526. doi:10.1161/STROKEAHA.115.012088
65. Rowe JB, Chan V, Ingemanson ML, Cramer SC, Wolbrecht ET, Reinkensmeyer DJ. Robotic assistance for training finger movement using a hebbian model: a randomized controlled trial. *Neurorehabil Neural Repair.* 2017;31:769-780. doi:10.1177/1545968317721975
66. Ingemanson ML, Rowe JR, Chan V, Wolbrecht ET, Reinkensmeyer DJ, Cramer SC. Somatosensory system integrity explains differences in treatment response after stroke. *Neurology.* 2019;92:e1098-e1108. doi:10.1212/WNL.0000000000007041
67. Reinkensmeyer DJ, Guigon E, Maier MA. A computational model of use-dependent motor recovery following stroke: optimizing corticospinal activations via reinforcement learning can explain residual capacity and other strength recovery dynamics. *Neural Netw.* 2012;29-30:60-69. doi:10.1016/j.neunet.2012.02.002
68. Norman SL, Lobo-Prat J, Reinkensmeyer DJ. How do strength and coordination recovery interact after stroke? A computational model for informing robotic training. *IEEE Int Conf Rehabil Robot.* 2017;2017:181-186. doi:10.1109/ICORR.2017.8009243
69. Milot MH, Spencer SJ, Chan V, et al. Corticospinal excitability as a predictor of functional gains at the affected upper limb following robotic training in chronic stroke survivors. *Neurorehabil Neural Repair.* 2014;28:819-827. doi:10.1177/1545968314527351