Supplemental Materials

Supplemental Methods

Kinematic data were represented as ($X_i(t), Y_i(t)$), where $X_i(t)$ and $Y_i(t)$ are the *X*- and *Y*-position of the hand at time *t*. FPCA expresses each movement as the combination of population-level components, selected to capture the major features of the kinematic data, and movement-specific weights or scores:

$$X_i(t) = \mu_X(t) + \sum_{k=1}^{3} c_{ik}^X \phi_k^X(t) \text{ and } Y_i(t) = \mu_Y(t) + \sum_{k=1}^{3} c_{ik}^Y \phi_k^Y(t) \quad (1)$$

Here $\mu_X(t)$ and $\mu_Y(t)$ are population mean functions, $\phi_k^X(t)$ and $\phi_k^Y(t)$ are shared components and the c_{ik}^X and c_{ik}^Y are the movement-specific scores. The mean and shared components were estimated using data from all movements and, given these, scores were estimated for each reaching trajectory. By construction, $\phi_k^X(t)$ and $\phi_k^Y(t)$ are data-driven characterizations of the major patterns observed in reaching trajectories. The scores c_{ik}^X and c_{ik}^Y quantify how these major patterns appear in each movement, and greatly reduce the dimension of the kinematic data: three *X* and *Y* scores suffice to explain more than 99% of the observed variance.

Subsequent analysis focused on these movement-specific scores. We computed the squared Mahalanobis distance ($MD_i^2 = (c_i - \bar{c})^T \Sigma^{-1} (c_i - \bar{c})$, where c_i is the concatenation of three *X* and *Y* scores and $\bar{c}$ is the element-wise mean of these vectors) as a measure of the distance of movement *i* from the population average. MD^2 was computed with respect to a reference population, which for

this study was a collection of reaching trajectories from the dominant arm of a group of naïve healthy age-matched controls with the same kinematic task. Intuitively, MD^2 is a vector-based analog to the squared Z-score in that it measures distance depending on the standard deviation of the reference population. Subject-specific average squared Mahalanobis distances (AMD^2) were computed for each subject at each target for each time point.

Supplemental Table 1. Participant clinical information

Participant	Stroke location	Lesion side	Edinburgh	Handedness	Star cancellation	Visual fields NIHSS
1	R MCA (frontal, temporal, BG, IC, CR)	Non-dominant	100	R	53 (no)	0
2	L MCA (BG,IC,CR)	Dominant	80	R	53 (no)	0
3	R MCA (frontal, parietal, temporal, insula)	Non-dominant	100	R	52 (no)	0
4	L MCA (BG,CR)	Dominant	100	R	52 (no)	0
5	R MCA (BG,CR)	Non-dominant	100	R	54 (no)	0
6	L MCA (IC)	Dominant	100	R	54 (no)	0
7	L pons + MCA (BG, frontal)	Dominant	80	R	52 (no)	1
8	R pons + CP + MCA (IC)	Non-dominant	80	R	52 (no)	0
9	R pons	Non-dominant	100	R	54 (no)	0
10	R MCA (CR)	Non-dominant	90	R	54 (no)	0
11	R MCA (CR, frontal, parietal)	Dominant	-100	L	54 (no)	0
12	R MCA (IC)	Non-dominant	100	R	51 (no)	0
13	R MCA (frontal)	Non-dominant	100	R	54 (no)	0
14	L MCA (CR, frontal, temporal)	Non-dominant	-40	L	51 (no)	1
15	L MCA (CR, frontal)	Non-dominant	-100	L	54 (no)	1
16	R MCA (IC)	Dominant	-100	L	54 (no)	0
17	R MCA (CR, frontal, parietal)	Non-dominant	100	R	54 (no)	0
18	R MCA (BG, IC, CR, temporal)	Non-dominant	90	R	47 (no)	0
	Abbreviations					
	L=Left					
	R=Right					
	MCA: Middle cerebral artery					
	IC: Internal Capsule					
	BG: Basal ganglia					
	CP: Cerebrabl Peduncle					
	CR: Corona Radiata					
	frontal, parietal, temporal, insula = ~ cortex					

Supplemental Results

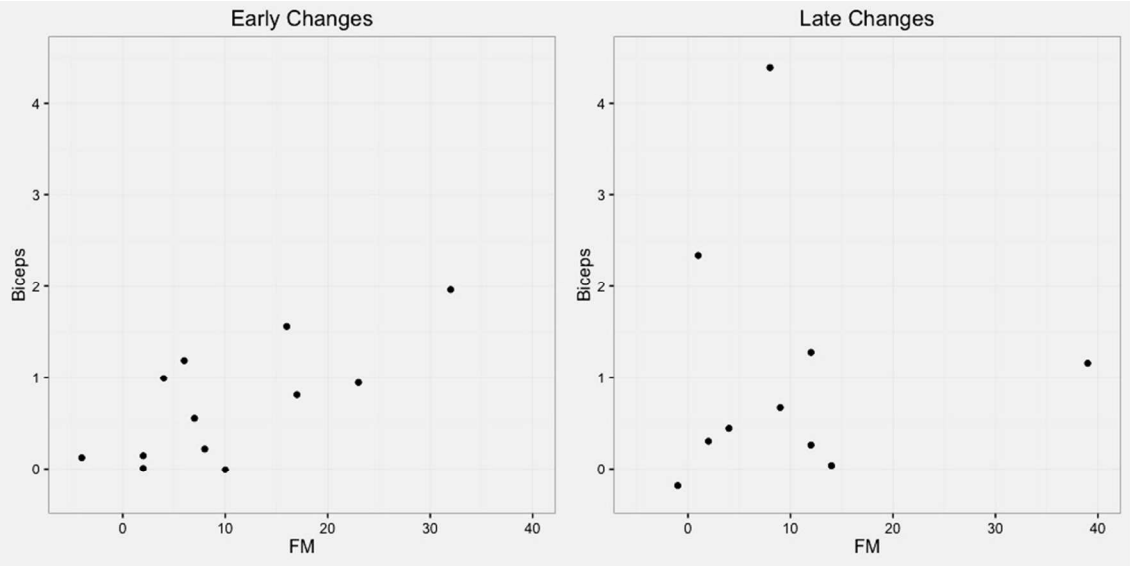
In Supplemental Fig 1, The left panel in the plot below shows the change in FMA-UE from Week 1.5 to Week 5 on the x-axis and the change in gross strength over the same period on the y-axis, with each point representing a subject. These data support the claim that early improvements in gross strength correlate with early improvements in FMA-UE (rho = .730). Similar observations can be made for the correlation between changes in ARAT and changes in gross strength (rho = .749).

The right panel shows analogous data for late recovery, from week 5 to week 54 which are complicated by two outliers who exhibit dramatic improvements in gross strength. These subjects were relatively fit and experienced only moderate initial impairment and, in addition to usual care, undertook a rigorous gross strength training program. Because their initial impairment was small late improvements in FMA-UE were limited; meanwhile, their gross strength improved markedly over the late recovery period.

With these subjects included, late improvements in gross strength do not correlate with late improvements in FMA-UE (rho = .023). However, omitting these subjects indicates that this correlation does exist for many subjects (rho = .622).

We believe these data are consistent with our fundamental hypothesis, but acknowledge that the existence of late correlations depends on decisions about subjects to be included in the analysis.

Meanwhile, on the complete dataset, we observe a larger decline in AMD^2 comparing Week 1.5 and 5 (delta = -25.10, $p < 0.001$) and no significant improvement in the later time period (delta = 0.30, $p = 0.93$). Meanwhile, the observed AMD^2 at Week 5 was similar using the complete data ($AMD^2 = 16.62$) and the data excluding “wrong” movements ($AMD^2 = 16.57$), providing more evidence for our claim that excluding motions primarily affects the magnitude of the deficit on the first visit without changing the interpretation of the results or the conclusions.



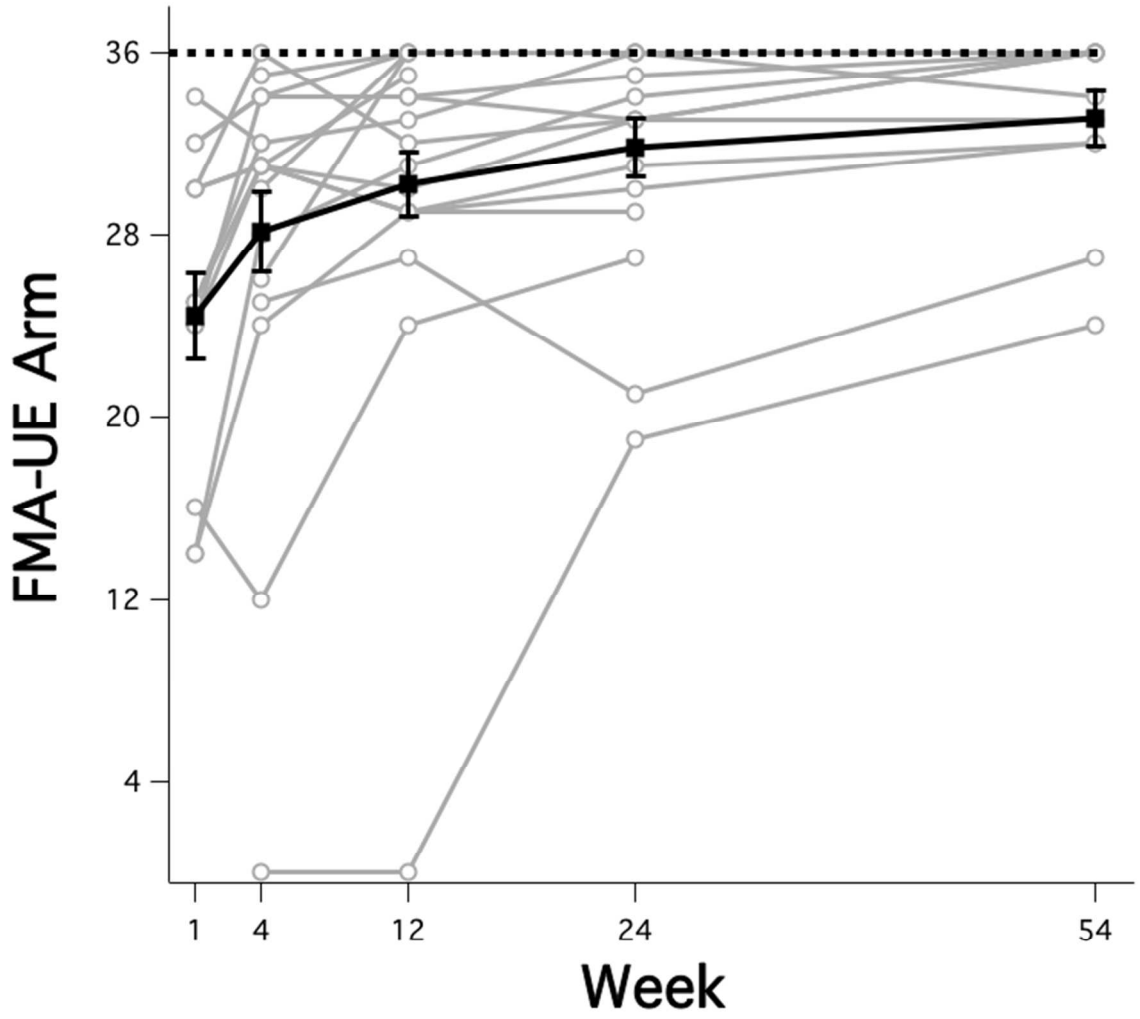
Supplemental Fig 1. Changes in Fugl-Meyer Assessment of Upper Extremity correlated with changes in biceps dynamometry, in A. Early recovery (week 0 to week 5) and B. Late recovery (week 5 to week 54)

Supplemental Table 2. Confidence intervals for all outcomes during early and late recovery.

	Early (Week 1.5 to Week 5)	**Late** (Week 5 to Week 54)
AMD[2]	-18.39 (-25.01, -11.76)	1.65 (-3.14, 6.44)
FMA-UE	8.23 (1.99, 14.47)	10.08 (5.36, 14.81)
ARAT	9.70 (1.58, 17.82)	9.98 (4.95, 15.02)
Biceps Dynamometry Z-score	0.57 (0.15, .99)	0.94 (0.32, 1.57)

Supplemental Table 3. P-values for all outcomes during early and late recovery adjusted for multiple comparisons.

	Early (Week 1.5 to Week 5)	**Late** (Week 5 to Week 54)
AMD[2]	-18.39 (adjusted $p < 0.001$)	1.65 (adjusted $p = 0.49$)
FMA-UE	8.23 (adjusted $p = 0.029$)	10.08 (adjusted $p = 0.002$)
ARAT	9.70 (adjusted $p = 0.038$)	9.98 (adjusted $p = 0.006$)
Biceps Dynamometry Z-score	0.57 (adjusted $p = 0.029$)	0.94 (adjusted $p = 0.015$)



Supplemental Figure 2. Proximal FMA-UE score. This score does not take into account wrist, hand, or coordination sections of the FMA-UE. FMA-UE proximal score for the arm keeps improving during the first year post-stroke. Early recovery delta for the group is 3.67 ($p = 0.034$) and late recovery delta is 5 ($p = 0.0003$)