**Additional information**

1. **Overview of the whole project**

Our project consists of four parallel RCTs targeting different symptoms and diseases (gait disturbance in stroke, speech disturbance in stroke, spinocerebellar degeneration, and parkinsonism) using different measures and interventions, with different target cortical regions and different imaginary tasks. These four projects were independently conducted, but only the feasibility measures (equipment related adverse effects and system errors) were pooled among these four RCTs. In addition, intervention methods (including the imagery tasks and target cortical areas) and endpoint measures were different.

We had planned 180 patients in total (40 for speech disturbance after stroke, 40 for parkinsonism, 40 for spinocerebellar degeneration (SCD), and 60 for gait disturbance after stroke). Finally, 117 patients (18 patients with Parkinsonism, 42 patients with SCD, and 57 patients with poststroke gait disturbance) were enrolled. Here, we report the results of RCT for gait disturbance after stroke.

1. **Rehabilitative intervention**

Patients in both groups were equally provided with customary rehabilitative intervention based on neurodevelopmental techniques for at least 5 days a week. The daily length of therapy was up to 180 minutes a day, including at least 60 minutes of physical therapy. Each therapy session included general conditioning; range-of-motion exercise for trunk and limbs; muscle strengthening; static and dynamic balance exercise with standing, kneeling, sitting, and quadruped standing; mobilizing the spine while prone and supine; walking indoors and outdoors; and climbing up and down stairs. The occupational therapy program emphasized improving activities of daily living and relaxation, hygiene, dressing, writing, eating, toileting, bathing, balance exercises, reaching, coordinative tasks of the upper limbs and trunk, and dual motor tasks such as handling objects while standing and walking. An additional 60 minutes of speech therapy was provided if a patient had problems with speech and swallowing. Duration for physical and occupational therapy in each patient was recorded from enrollment to final clinical assessment.

1. **Outcome Measures**

*3m-Timed Up and Go test (TUG)*

The TUG measures the time in which the patient rises from a chair, walks 3 meters, turns, walks back, and sits down again.1 In the study, participants sat on a chair without using back support, and time was measured from the moment when their buttocks lifted from the chair to the moment when their buttocks touched the seat again. Time was recorded by a stopwatch. Each participant performed TUG with two different turning directions, and the best performance was adopted.

*Berg Balance Scale (BBS)*

The BBS is a functional balance measurement and consists of 14 items on a 5-point ordinal scale ranging from 0 to 4.2 In this scale, 0 indicates unable to perform the task and 4 indicates able to complete the task. Total scores can range from 0 to 56 with higher scores indicating better postural ability.

*Functional Independence Measure (FIM)*

FIM is an assessment tool for evaluating functional independence of patients in activities of daily living after stroke.3 It consists of 18 items ranging from 1 indicating completely dependent to 7 indicating independent without device. The FIM includes 13 items assessing motor aspect of daily functioning, and 5 items assessing cognitive aspects of daily functioning. In total, FIM scores can range from 18 to 126 with higher scores indicating better daily function.

*Gait Speed*

Participants are instructed to walk straight as fast as possible, for 13m. If participants could not walk independently at baseline, walking aids such as canes or orthosis were used, and the same walking aids were subsequently also used at follow up assessment. The time in the middle 10m was recorded by a stopwatch.

*Fugl-Meyer Assessment (F-M)*

The F-M is a stroke-specific, performance-based impairment index to assess motor functioning after stroke. It examines movement, coordination, and reflex action of the upper and lower extremities.4 For upper limb motor functioning, the examiner assesses 31 items of a 3-point ordinal scale ranging from 0 to 2 and 2 items of a 2-point ordinal scale of 0 or 2. In total, F-M for upper extremities can range from 0 to 66, with higher scores indicating better motor function. For lower limb motor function, the examiner assesses 15 items of a 3-point ordinal scale ranging from 0 to 2 and 2 items of a 2-point ordinal scale of 0 or 2. In total, F-M for lower extremities can range from 0 to 34, with higher scores indicating better motor function.

1. **Allocation of the patients**

In this study, a computer-generated sealed envelope method was adopted for patient allocation to REAL- and SHAM-group. First, one author who was not involved in treatment nor clinical and imaging evaluation (KK) ran the randomization program to generate the allocation order. In this program, the value of “0” or “1” was given for each registration ID (from 1 to 80), under the condition that each consecutive eight ID block (1~8, 9~16 etc.) has four “1s”. This “registration order” data was embedded in the neurofeedback program, and a copy of the data was held by KK, and key-opened at the time of interim and final analysis. After interim analysis, the “registration order” data was reformed and re-embedded for further intervention. In this study, each patient provided unique registration ID (according to the registration order), and the neurofeedback program used this ID and “registration order” data to allocate to the REAL and SHAM groups. If “registration order” data of ID number X is “1”, the neurofeedback program used the online-provided fNIRS data (data from the participant who is actively performing the task) for calculating the cortical activation. If “registration order” data is “0”, the neurofeedback program used the pre-recorded fNIRS data (other subjects’ data).

1. **Neurofeedback intervention**

During neurofeedback intervention, patients sat comfortably in an armchair with a headrest and their arms on the armrests. At the start of each neurofeedback training, they watched 10 minutes of video instructions on the kinesthetic motor imagery task, which consist of rising up from the chair followed by a step twice task and walking along the corridor task. Neurofeedback-based training was performed using the fNIRS mediated neurofeedback system, 5, 6 and fNIRS apparatus was attached to patients during video-based instruction. After instruction, patients were asked to perform kinesthetic motor imagery of both tasks with supplementary motor area (SMA) activation feedback for 10 minutes. Patients were asked to open their eyes to see the feedback signal (horizontal bar) on the screen. Participants were instructed that the bar represents the contrast of cortical activation between the motor imagery task and the rest, therefore, continuously increased bars represent efficacious motor imagery. Patients were asked to try to keep the bar as high as possible. The feedback task was comprised of 16 repetitions of a 5-s trial, with pseudo-randomized rest periods ranging from 8 to 16s between each task period. The start and end of the trial were signaled with audio cues (start, single beep; end, double beep). Although threshold values were not presented to the participants, the color change in the feedback bar was generally recognizable when a t-value of 2 was attained. Examiners visually inspected whether patients moved their trunk or lower-limb during motor imagery task. Patients in both groups performed the exact same procedures for the neurofeedback training, and examiners, patients, and assessors did not know any information about the group allocation (triple-blinded).

1. **Image acquisition and analysis**

*fNIRS measurement and detector setting*

To detect motor-imagery related cortical hemodynamic changes, we used a continuous wave fNIRS system with 16 light sources and detectors (OMM-3000; Shimadzu Corp., Kyoto, Japan). In this system, three different wavelengths of infrared light at 780, 805, and 830 nm were applied, based on a modified Beer-Lambert law,7 to calculate relative concentration change of oxygenated hemoglobin (OxyHb) and deoxygenated hemoglobin (DeoxyHb) using an arbitrary unit (mmol/L × mm). Optodes were placed on the fronto-parietal scalp using a custom made hard-plastic holder with a regular inter-optode distance of 3 cm. To eliminate extra-brain contamination including scalp blood-flow,8 four short-distance (1.2 cm) channels were added on the bilateral prefrontal scalp. The light source at the center of the third row was placed at the vertex (Cz) of the subject, and fNIRS channel was defined as the midpoint of each corresponding light source-detector pair. Cortical activity was measured from 52 channels at 4 Hz. Similar to our previous studies,5, 6, 9 we estimated the position of each channel using individual anatomical 3D T1-weighted magnetic resonance (MR) images from 36 participants, for whom MR images could be obtained. The spatial configuration of the optodes on the scalp was estimated using a virtual holder set, and normalized to the standard ICBM152 (ICBM, http://www.loni.ucla.edu/ICBM/) template10 using SPM8 software (<http://www.fil.ion.ucl.ac.uk/spm/>). The location of each optode was also estimated on the Montreal Neurological Institute (MNI) standardized scalp10 via an affine transformation matrix. The cortical projection point of each channel was estimated using the balloon-inflation method,11 and the averaged coordinates and their variance across patients were estimated. Since the dispersion of the estimated channel positions was within several millimeters, we considered that each channel location was similar across patients. The cortical region covered by each channel was estimated using MRIcro software (http://www.MRIcro.com), which provided Brodmann areas and automated anatomical labeling.12

Cortical registration results confirmed that the inter-subject variability of the fNIRS channels were within a few millimeters, and we considered that each channel covers similar cortical regions in all the patients (Figure S1). At the first session, patients in SHAM feedback group showed lateral premotor activation in unaffected hemisphere during imagery of rising-up from the chair, but no significant cortical activation during imagery of walking along. Patients in REAL feedback group showed no significant cortical activation during imagery of rising-up from the chair but they showed significant activation in the affected prefrontal and frontal eye field activation, and unaffected sensorimotor and parietal cortex during imagery of walking along. At the last session, patients in SHAM feedback group showed significant cortical activation during both imagery tasks, whereas patients in REAL feedback group showed wide-spread cortical activation including bilateral SMA, premotor cortex (PM), and prefrontal cortex during two different gait and balance related motor imagery tasks.

1. **FNIRS data analysis**

Since we included patients with either a right- or left-sided lesion, all imaging data from the patients with a left-sided lesion were flipped horizontally before data analysis, to arrange the affected hemisphere as right-sided. For neurofeedback, task-related signal changes were estimated by comparing β-coefficients between task and resting data using one-sided (right-tailed) paired t-tests in a real-time fashion using adaptive GLM-based analysis using data from 20 s sliding time-windows.5, 6, 9 The calculated t-values were used as markers of cortical activation at each channel. The largest calculated t-values for the four channels covering the SMA (21, 22, 28, and 29) was shown as the height and color of the vertical feedback bar, to provide feedback for the participants. Averaged feedback signal values for each patient was recorded.

For post-hoc analysis of motor imagery related cortical activation changes, fNIRS data was analyzed using a general linear model using an in-house program running on MATLAB.13 The preprocessing procedure included removing baseline drift with a high-pass filter (cut-off frequency = 0.03 Hz). To investigate the motor imagery related cortical activation before and after neurofeedback, we performed channel-based estimation for each subject using first and last (6th) neurofeedback session data. In addition, to estimate the effect of neurofeedback on cortical activity, we made channel-based within-subject contrast images, comparing data from last and first neurofeedback session. Next, we performed a second-level group analysis adopting a random-effects model with individual contrasts used as the dataset (Figure S2). For each condition (REAL and SHAM), two-tailed one-sample t-tests were performed against a mean of zero, to detect imagery-related cortical activation and its changes by means of neurofeedback. Finally, we performed two-tailed unpaired t-tests comparing the neurofeedback effect on imagery-related cortical activation between two feedback conditions (REAL versus SHAM). In the channels showing significant imagery-related cortical activation change, we also performed channel-based correlation analysis between gait and balance improvement (TUG time, BBS score, and speed change) and cortical activation changes with Pearson’s correlation analysis. Statistical significance was set at q < 0.05 (FDR-correction was performed for multi-channel recording of cortical activation).

1. **MRI measurement**

In Morinomiya Hospital, we obtained spin-echo echo-planar imaging under the following conditions; time points = 240, TR = 2500 ms, TE = 30 ms, flip angle = 80°, matrix = 64 × 64, FOV = 220 mm, slice thickness = 3.5 mm and slice number = 40. For rsfMRI scans, we instructed participants to lay still with their eyes open and to look at the fixation cross without thinking about anything for 10 minutes. Anatomical scans were used for registration; a sagittal 3D fast-spoiled gradient-recalled echo-pulse sequence was used with the following parameters: echo time (TE) = 2.7 ms, repetition time (TR) = 7.0 ms, inversion time = 400 ms, matrix dimensions = 256 × 256, field of view (FOV) = 240 mm, slice thickness = 1.2 mm and slice number = 200. Lesion location was confirmed by the anatomical scans and individual lesion masks were manually generated for rsfMRI analysis and lesion mapping. All MRI data were acquired using a Philips 1.5T Intera Achieva Nova (Philips Medical Systems, Best, the Netherlands).

In Osaka University, we obtained spin-echo echo planar imaging for rsfMRI scan under the following conditions; time points = 240, TR = 2500 ms, TE = 30 ms, flip angle = 80°, matrix = 64 × 64, FOV = 220 mm, slice thickness = 3.5 mm and slice number = 40. During rsfMRI scan, we instructed participants to lay still with their eyes open and to look at the fixation cross without thinking about anything for 10 minutes. Anatomical scans were used for registration; a sagittal 3D fast-spoiled gradient-recalled echo-pulse sequence was used with the following parameters: echo time (TE) = 2.7 ms, repetition time (TR) = 7.0 ms, inversion time = 400 ms, matrix dimensions = 256 × 256, field of view (FOV) = 240 mm, slice thickness = 1.2 mm and slice number = 200. All images were acquired by using a 3-T MR scanner (GE Healthcare, Milwaukee, WI, USA).

1. **MRI data analysis**

As stated above (7. MRI measurement), we used a different protocol and MRI apparatus in the two study sites. Because the magnetic power and developer of each MR apparatus is different, we assumed that it was not appropriate to analyze data from different sites together. Therefore, we used data from one study site where more patients’ data were obtained (Morinomiya Hospital). We obtained structural MRI data for fNIRS channel registration, and resting state functional MRI data, to investigate whether neurofeedback can modulate resting state functional network. All MRI data from the patients with a left-sided lesion were flipped horizontally before data analysis in this study (lesion overlap is presented in Figure 3A). We investigated the longitudinal changes of resting-state functional connectivity from the SMA, which is the target cortical area for neurofeedback facilitation. After preprocessing procedure including normalization, segmentation, smoothing, de-noising, and band-pass filtering (0.01–0.1 Hz) of structural and functional MRI images, we first evaluated whole-brain functional connectivity from bilateral SMA. All analyses procedures were performed with the CONN-fMRI Functional Connectivity toolbox v17 (http://www.nitrc.org/projects/conn).14 We investigated group differences of the SMA-related resting-state functional connectivity changes and the correlation between the SMA-related functional connectivity changes and recovery of clinical measures. For connectivity analysis, a threshold of uncorrected *p* < 0.001 with a false discovery rate-corrected *q* < 0.05 at the cluster level was considered statistically significant.

1. **Sample size estimation**

Sample size estimation was conducted using G\*Power version 3.1 (<http://www.gpower.hhu.de/>). Previously, we have conducted a pilot study investigating the NIRS neurofeedback effect on post-stroke upper limb paresis,5 which included 10 patients in the intervention and control groups. In this pilot study, the average Fugl-Meyer Scale change, which was the primary outcome, was 4.7 and 2.3 for intervention and control group respectively (pooled SD = 2.46). Using this data, we calculated the effect size for a two group comparison (d) as 0.97, and calculated a minimum sample size as 36 (18 patients for each groups) with two-tailed α of 0.05 and a power (1-β) of 0.80, for the comparison of 2 independent group differences for patients with upper limb impairment and speech disturbance. Considering that the relatively large inter-subject variability of gait analysis measures such as TUG or gait speed, we estimate that the pooled SD for the patients with gait and balance disturbance would be larger than that of patients with upper limb paresis. Therefore, we estimated a 15% larger SD in effect size estimation, resulting in an effect size (d) of 0.85. Using this data, we estimated the minimum sample size required to reach 80% power as 46 (23 for each group). To investigate the group comparisons of primary and secondary measures, data from patients who completed the intervention was used. All records from recruited patients were used for investigating the safety measures.

**Figure S1:** fNIRS channel overlap in all subjects

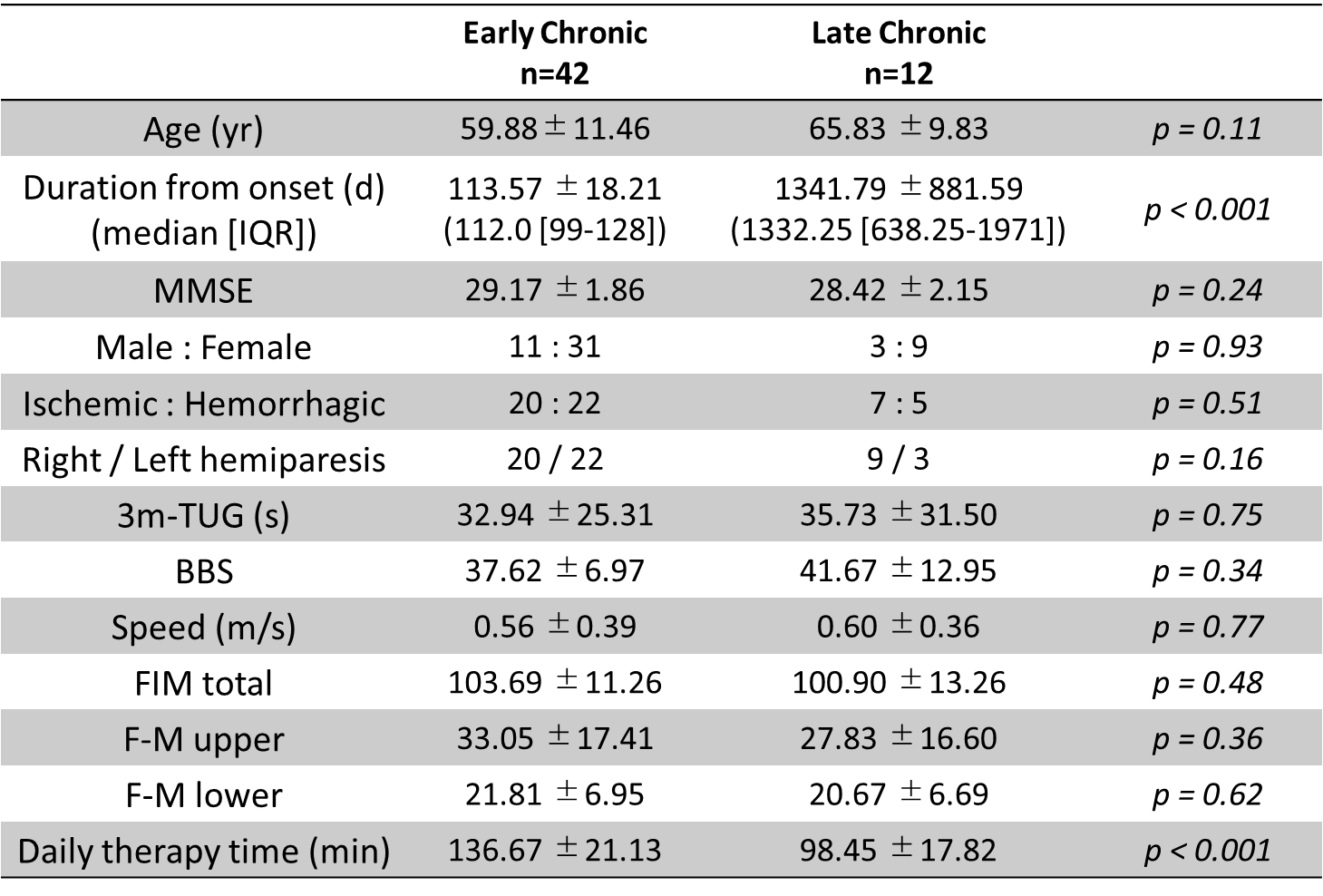
(A) Green dots indicate estimated channels in each subject, and dark blue dots indicate averaged estimated cortical point covered by each fNIRS channel. The colored area represents the range within mean +1 SD of the averaged distance from the averaged cortical point to each subject’s estimated points. (B) fNIRS channels cover the fronto-parietal scalp and the signals from these channels are suspected to cover the SMA (Ch. 21, 22, 28, and 29). These were used for real-time processing and in the calculation of the feedback signal.

**Figure S2:** Gait and balance related cortical activation changes during neurofeedback intervention in both groups.

At the first session, there was no group difference in imagery related cortical activation. However, at the last session, gait imagery related cortical activation was significantly more facilitated in the REAL group.

**Supplementary Video:** Representative gait performance in patient from the REAL feedback group (recorded in April 2014 at Osaka University Hospital). Gait speed of the patient is improved after two weeks’ neurofeedback intervention.

**Table S1.** Baseline characteristics of patients in both early (within 150 days from onset) and late (after 150 days from onset) intervention groups.

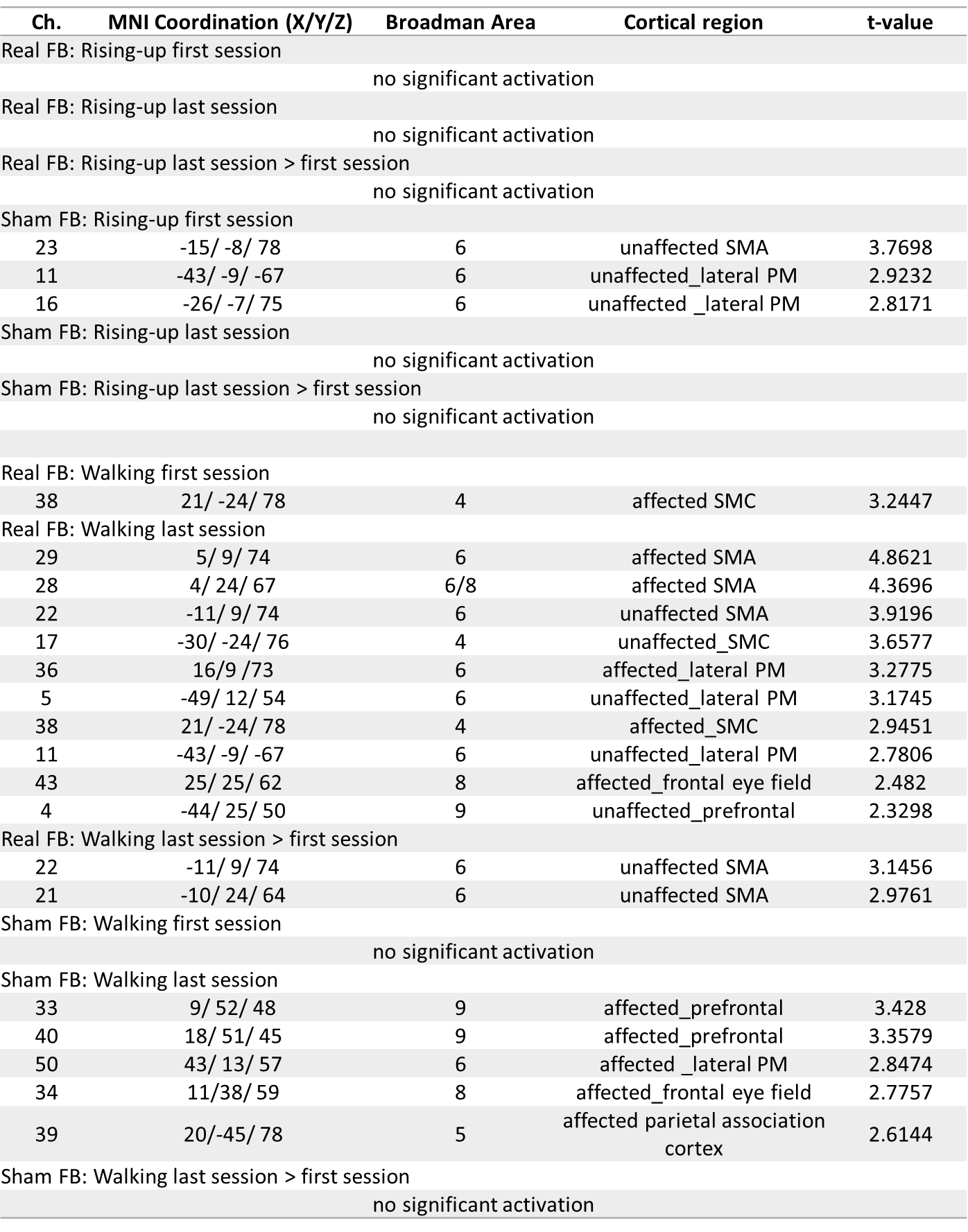


Data are presented as Mean ± Standard deviation, or ratios.

MMSE: Mini-mental State Examination, 3m-TUG: 3 meter-Timed Up and Go test, BBS: Berg Balance Scale, FIM: Functional Independence Measure, F-M: Fugl-Meyer motor assessment scale

MMSE: Mini-mental State Examination, 3m-TUG: 3 meter- Timed Up and Go test, BBS: Berg Balance Scale, FIM: Functinoal Independence Measure, F-M: Fugl-Meyer motor assessment scale

**Table S2.** Gait and balance related cortical activation before and after neurofeedback intervention in both groups.



Ch: Channels in fNIRS recording, MNI: Montreal Neurological Institute, Real FB: REAL feedback group, Sham FB: SHAM feedback group

**References**

1. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39:142-148.

2. Berg K, Wood-Dauphinee S, Williams JI. The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. Scand J Rehabil Med 1995;27:27-36.

3. Linacre JM, Heinemann AW, Wright BD, Granger CV, Hamilton BB. The structure and stability of the Functional Independence Measure. Arch Phys Med Rehabil 1994;75:127-132.

4. Fugl-Meyer AR, Jaasko 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.

5. Fujimoto H, Mihara M, Hattori N, et al. Cortical changes underlying balance recovery in patients with hemiplegic stroke. Neuroimage 2014;85 Pt 1:547-554.

6. Mihara M, Miyai I, Hattori N, et al. Neurofeedback using real-time near-infrared spectroscopy enhances motor imagery related cortical activation. PLoS One 2012;7:e32234.

7. Cope M, Delpy DT, Reynolds EO, Wray S, Wyatt J, van der Zee P. Methods of quantitating cerebral near infrared spectroscopy data. Adv Exp Med Biol 1988;222:183-189.

8. Takahashi T, Takikawa Y, Kawagoe R, Shibuya S, Iwano T, Kitazawa S. Influence of skin blood flow on near-infrared spectroscopy signals measured on the forehead during a verbal fluency task. Neuroimage 2011;57:991-1002.

9. Fujimoto H, Mihara M, Hattori N, et al. Neurofeedback-induced facilitation of the supplementary motor area affects postural stability. Neurophotonics 2017;4:045003.

10. Mazziotta J, Toga A, Evans A, et al. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos Trans R Soc Lond B Biol Sci 2001;356:1293-1322.

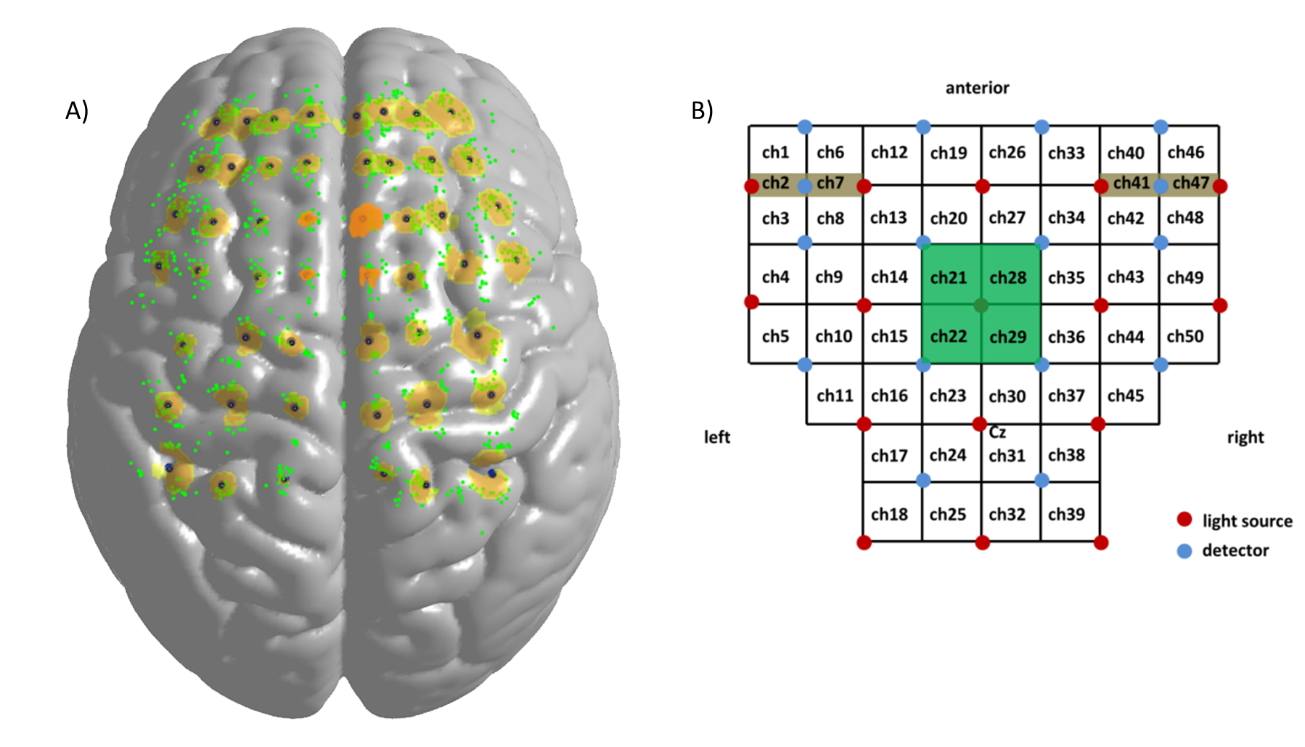
11. Okamoto M, Dan I. Automated cortical projection of head-surface locations for transcranial functional brain mapping. Neuroimage 2005;26:18-28.

12. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 2002;15:273-289.

13. Mihara M, Miyai I, Hatakenaka M, Kubota K, Sakoda S. Role of the prefrontal cortex in human balance control. Neuroimage 2008;43:329-336.

14. Whitfield-Gabrieli S, Nieto-Castanon A. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. Brain Connect 2012;2:125-141.

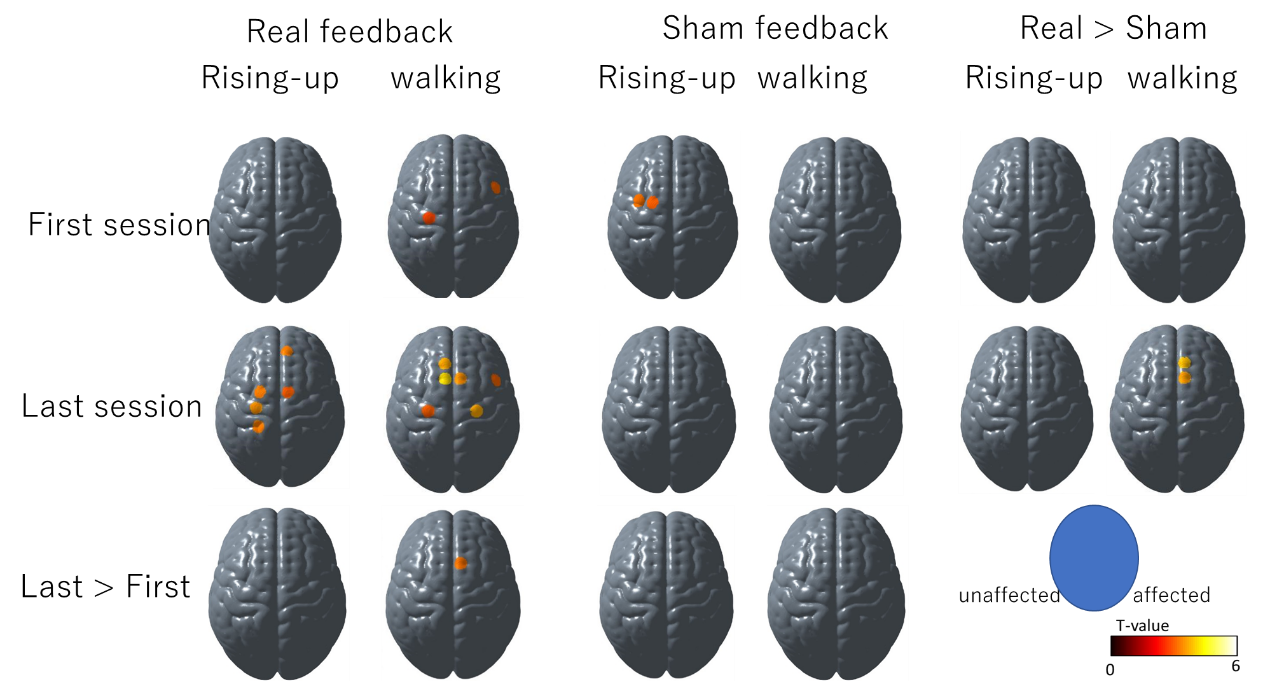
**Figure S1**



**Figure S1:** fNIRS channel overlap in all subjects.

(A) Green dots indicate estimated channels in each subject, and dark blue dots indicate averaged estimated cortical point covered by each fNIRS channel. The colored area represents the range within mean +1 SD of the averaged distance from the averaged cortical point to each subject’s estimated points. (B) fNIRS channels cover the fronto-parietal scalp and the signals from these channels are suspected to cover the SMA (Ch. 21, 22, 28, and 29). These were used for real-time processing and in the calculation of the feedback signal.

**Figure S2**



**Figure S2:** Gait and balance related cortical activation changes during neurofeedback intervention in both groups.

At the first session, there was no group difference in imagery related cortical activation. However, at the last session, gait imagery related cortical activation was significantly more facilitated in the REAL group.