**Supplementary Material**

**SUPPLEMENTARY METHODS**

*Sample size*

 The lack of prior studies with similar sample population, intervention, and outcome measures did not allow an a priori power analysis. Sample size was a priori determined taking into account sample sizes from previous studies and reviews on cognitive rehabilitation, as well as feasibility and timing, and a foreseen target of 40 patients was established. According to the study chronogram, administration of the neuropsychological rehabilitation program lasted 18 months, and 20 treated patients corresponded to a total of 800 h of individual attention process training, and to approximately 45 h of rehabilitation activities in a month. This workload was in line with the availability of dedicated space, setting, and personnel.

*Randomization*

 The RehAtt study included an intervention group (‘attention training’) and a control group (‘standard care’). After baseline assessment, participants were randomly assigned to ‘attention training’ or ‘standard care’ group. Stratified randomization was used to ensure the balance for possible prognostic factors (age, gender, and WMH degree) across the groups.

*Interventions*

 Participants in the ‘attention training’ group were scheduled to have weekly sessions of 2 h of individual attention process training for a total of 20 weeks. The APT-II program consists of a group of hierarchically organized tasks that exercise different components of attention, including: focused, sustained, selective, alternating, and divided attention. Some tasks are subject-paced while others are experimenter-paced, but in both cases, exercises typically take 2-3 min to complete. The program tasks place increasing demands on complex attentional control and working memory systems, but do not have a heavy memory load and usually require the participant to classify stimuli. Examples of exercises include auditory attention tapes such as listening to descending number sequences, alphabetizing words in an orally presented sentence, detecting targets with the presence of distracter noise or complex semantic categorization tasks requiring switching sets. A number of tasks combine auditory and visual activities. Feedback about accuracy and speed of performance is provided after each exercise. Other aspects of performance, such as patterns of errors, are also discussed. As the program proceeds, participants are educated about different types of attention and parallels between difficulties of daily living and problems performing particular APT-II exercises are pointed out.

 Participants in the ’standard care’ group did not receive cognitive interventions nor an active placebo treatment. We choose not to engage the control group in an alternative treatment because it could reduce the patients’ compliance in the study, be very demanding, and, most importantly, questionable under ethics viewpoint. Participants in the ’standard care’ group have been instructed to have a usual lifestyle and to immediately communicate any change regarding medications or activities focused on cognition.

*Outcome measures*

Functional and quality of life scales (primary outcomes)

* Activities of Daily Living scale (ADL): a measure of participant's skill level on six basic functional domains. We took into account the number of preserved items and summed them into a global score ranging from 0 (completely dependent) to 6 (completely autonomous).
* Instrumental Activities of Daily Living scale (IADL): a measure of participant's skill level on eight complex functional domains. We took into account the number of impaired items, and summed them into a global score ranging from 0 (completely autonomous) to 8 (completely dependent).
* Disability Assessment in Dementia scale: a measure of participant's skill level on several basic, instrumental and leisure activities. The Disability Assessment in Dementia scale includes 40 dichotomous items summed into a total score then converted into a percentage (higher scores represent less disability).
* Short Form Health Survey (SF-36): a 36-item self-reported survey of patients’ own general health and well-being. The SF-36 consists of eight basic health dimensions summarized in 2 higher-order summary scores (Physical Component Summary and Mental Component Summary). Summary scores are transformed into a 0-100 scale (lower scores represent more disability), and t scores are calculated using national norms.
* EuroQol: a self-reported health-related quality of life measure composed of a descriptive section converted into a single summary index, and a visual analogue scale used as a quantitative measure ranging from 0 (worst imaginable health state) to 100 (best imaginable health state).
* Attention Questionnaire: a 12-item self-reported questionnaire specifically developed to rate the occurrence of different attention problems in everyday life. The Attention Questionnaire supplies a total score ranging from 0 (absence of attention problems) to 36 (highest presence of attention problems).
* Geriatric Depression Scale: is a 15-item [self-report](https://en.wikipedia.org/wiki/Self-report_inventory)ed questionnaire used to identify depressive symptoms in the elderly. The Geriatric Depression Scale supplies a total score ranging from 0 (absence of depressive symptoms) to 15 (severe depressive symptoms).

The VMCI-Tuscany neuropsychological battery (secondary cognitive outcomes)

* Global cognitive functioning: Montreal Cognitive Assessment test (MoCA), and Mini Mental Status Examination test (MMSE).
* Memory: Rey Auditory-Verbal Learning Test (RAVL) immediate and recall, Short story, and Rey–Osterrieth Complex Figure (ROCF) recall.
* Attention and executive functions: Trail Making Test (TMT) part A and B, Visual search, Symbol Digit Modalities Test (SDMT), and Color Word Stroop Test.
* Language: phonemic and semantic verbal fluency.
* Constructional praxis: Rey–Osterrieth Complex Figure (ROCF) immediate copy.

 Cognitive performance was evaluated based on the non-parametric equivalent score methodology for all the tests except for the Symbol Digit Modalities Test for which the national parametric norms were used because equivalent scores were not available.

*MRI examination*

 All patients were examined on a clinical 1.5 T scanner (Aera, Siemens Medical Solutions, Erlangen, Germany) with 46 mT/m maximum gradients strength, 200 mT/m/ms slew-rate and 20-channel radiofrequency head and neck coil. After scouts, the examination protocol included high-resolution sagittal contiguous 3D T1-weighted images with 1-mm isotropic voxels which were obtained with a Turbo Spin Echo SPACE (Sampling PErfection with Application optimized Contrasts using different flip angle Evolution) sequence [repetition time (TR)=600 ms, echo time (TE)=7 ms, slice thickness=1 mm, no inter-slice gap, field of view (FOV)=232 mm (phase) × 256 mm (frequency), acquisition matrix=232×256, turbo factor=36, number of excitations (NEX)=1, GRAPPA acceleration factor = 2, number of slices = 176]. The rsfMRI data were obtained using a T2\*-sensitive EPI sequence (TR=2.520 s, TE=50 ms, flip angle=90°, field of view=256 mm × 256 mm, acquisition matrix=64 × 64, GRAPPA acceleration factor=2, no inter-slice gap, number of slices=32). Two hundred and fifty volumes were acquired for a total acquisition time of 10.5 min. During rsfMRI acquisition, the subjects were instructed to lie still with their eyes closed and not to think of anything in particular. Cushions were used to minimize head motion during the scan.

*RsfMRI pre-processing*

 RsfMRIdata pre-processing was carried-out using FSL (FMRIB's Software Library, http://www.fmrib.ox.ac.uk/fsl). The following pre-statistics processing was applied: removal of the first 5 volumes of each scan to avoid T1-related relaxation effects; motion correction using MCFLIRT (Motion Correction FMRIB's Linear Image Registration Tool) [1]; slice-timing correction using Fourier-space time-series phase-shifting; grand-mean intensity normalization of the entire 4D dataset by a single multiplicative factor; low-pass filtering at 0.08 Hz of the time series of rsfMRI data of each voxel [2]. To control for motion artifacts the mean (across voxels) voxel absolute displacement (each time point with respect to the reference image, i.e. to the middle-time point rsfMRI image) was calculated by MCFLIRT in each patient. Two patients had one of the two (baseline or follow-up) rsfMRI examinations with mean voxel absolute displacement greater than 2 mm and were thus excluded.

*Computation of ReHo maps*

 Co-registration of the reference rsfMRI images to the individual high resolution T1-weighted image was carried out using affine registration implemented in the ANTs package [3]. The individual high-resolution T1-weighted images were co-registered to the standard-space Montreal Neurological Institute (MNI) 152 brain using an affine followed by a non-linear transformation using the ANTs package [3]. Regional homogeneity (ReHo) was calculated in the native rsfMRI space using the Data Processing Assistant for Resting-State fMRI (http://rfmri.org/DPARSF) [4]. ReHo maps were computed using the Kendall coefficient of concordance [5] within a 27-voxel cubic neighborhood. For a given voxel, a high ReHo value implies that, within the 27-voxel neighborhood, resting-state time series have high synchronization with each other. For each patient, a brain mask was also obtained from the reference rsfMRI image using the brain extraction tool (BET), part of FSL (FMRIB’s Software Library). Each individual ReHo map was transformed into a standard Z-value map by subtracting the global mean and dividing by the standard deviation within the brain mask, and normalized to the MNI standard space using the transformation previously computed when co-registering rsfMRI images to the MNI standard space.

*Statistical analysis of rsfMRI data*

 The p-values were calculated employing permutation-based statistics (5,000 permutations) and corrected for multiple comparisons using the “3D” parameter settings with threshold-free cluster enhancement (TFCE), thereby avoiding the use of an arbitrary threshold for the initial cluster formation [6]. A p-value <0.05 corrected for multiple comparisons across space (family-wise error rate correction) was considered statistically significant [7-8]. The between-group analysis was performed in the MNI152 space. In order to anatomically map the significant clusters of synchronization identified at the group analyses, we overlaid the corresponding statistic maps onto the Cerebellar Atlas in MNI152 space after normalization with FNIRT [9].

**Supplementary Table 1.** Extent and MNI coordinates of the clusters where the difference of Z-transformed ReHo between follow-up and baseline is greater in the attention training group as compared to the standard care group (TFCE, p<0.05).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Cluster****number** | **Cluster extent (mm3)** | **Max X****(mm)** | **Max Y****(mm)** | **Max Z****(mm)** | **Anatomical area** |
| 1 | 2184 | -4 | -62 | -34 | Vermis VIIIb |
| 2 | 848 | 22 | -68 | -46 | Right VIIb lobule |
| 3 | 568 | -28 | -68 | -46 | Left VIIb lobule |

Coordinates are expressed in MNI152 standard space.

Anatomical area refers to the most probable area according to the Cerebellar Atlas in MNI152 space after normalization with FNIRT.

 **REFERENCES**

1. Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825-841.
2. Lv Y, Margulies DS, Villringer A, Zang YF (2013) Effects of finger tapping frequency on regional homogeneity of sensorimotor cortex. *PLoS One* **8**, e64115.
3. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC (2011) A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* **54**, 2033-2044.
4. Chao-Gan Y, Yu-Feng Z (2010) DPARSF: A MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Front Syst Neurosci* **4**, 13.
5. Zang Y, Jiang T, Lu Y, He Y, Tian L (2004) Regional homogeneity approach to fMRI data analysis. *Neuroimage* **22**, 394-400.
6. Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* **44**, 83-98.
7. Holmes AP, Blair RC, Watson JD, Ford I (1996) Nonparametric analysis of statistic images from functional mapping experiments. *J Cereb Blood Flow Metab* **16**, 7-22.
8. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE (2014) Permutation inference for the general linear model. *Neuroimage* **92**, 381-397.
9. Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N (2009) A probabilistic MR atlas of the human cerebellum. *Neuroimage* **46**, 39-46.