Supplementary Material

The Mini-Mental State Examination primarily detects cognitive impairment due to left middle cerebral artery infarcts

Supplementary Methods

Magnetic resonance imaging

A 3.0 tesla MRI scanner (Achieva, Philips Healthcare, Eindhoven, the Netherlands) was used in both hospitals. The Hallym VCI cohort MRI protocol consisted of DWI, axial T1- and T2-weighted spin echo, FLAIR, gradient-echo imaging, and coronal T2-weighted spin echo imaging. The following acquisition parameters were set for FLAIR images: repetition time, 11,000 ms; echo time, 125 ms; inversion time, 2,800 ms; slice thickness, 5 mm; intersection gap, 2 mm; matrix, 512 x 512; flip angle 90 degree. The acquisition parameters for DWI images were: repetition time, 3,000 ms; echo time, 56 ms; diffusion b-value 1,000; slice thickness, 5 mm; intersection gap, 0 mm; matrix, 256 x 256; flip angle 90 degree. Patients in the Hallym VCI cohort were scanned in the first week of hospital admission. The Bundang VCI cohort MRI protocol consisted of DWI, axial T1- and T2-weighted spin echo, FLAIR, gradient-echo imaging, and coronal T1-weighted spin echo imaging. FLAIR imaging was obtained using a fast-spin echo sequence. The following imaging parameters were set for FLAIR images: repetition time, 11,000 ms; echo time, 125 ms; inversion time, 2,800 ms; slice thickness, 5 mm; intersection gap, 1 mm; matrix, 512 x 512; flip angle 90 degree. DWI imaging was obtained using an EPI-spin echo sequence. The acquisition parameters for DWI images were set as follows: repetition time, 5,000 ms; echo time, 50 ms; diffusion b-value 1,000; slice thickness, 5 mm; intersection gap, 1 mm; matrix, 256 x 256; flip angle 90 degree. Patients in the Bundang VCI cohort were scanned at hospital admission and at one-week follow-up; the follow-up scan was used for the current study.
Generation of lesion maps

Infarct segmentations were performed by two trained investigators (A.K.K. and G.A.). They were subsequently checked and adapted by an experienced rater (N.A.W.), and further revised by another experienced rater (J.M.B.) in the event of uncertainty regarding lesion location or classification. Discrepancies between ratings were discussed in consensus meetings. All four raters were blinded to the neuropsychological data during the segmentation process. The raters had access to all available MRI sequences and the time interval between stroke onset and MRI. In the acute stage (≤2 weeks after stroke onset), DWI was used as a preferred modality for segmentation because it is considered the optimal sequence (i.e., is generally the most sensitive sequence and provides high contrast between infarcted and normal brain tissue) to visualize acute infarcts within this time frame.[1] If DWI was not available, or a comparison between the FLAIR and DWI revealed that the FLAIR provided a more accurate image of the infarct, or MRI was performed in a more chronic stage (i.e., in 10 patients MRI was performed >2 weeks after stroke), infarcts were segmented on FLAIR sequences (N=30; 2.5%). ADC and T1-weighted sequences were used as a reference to support identification and delineation of the infarcts.

Intra-observer (two ratings by N.A.W. with a one-year time interval) and inter-observer (between A.K.K./G.A. and N.A.W.) agreement were determined using the Dice Similarity Coefficient (DSC),[2,3] which expresses the degree of overlap between the segmented voxels on a scale from 0 to 1, on a random subset of 30 DWI scans. Both intra-observer (DSC=0.87; SD=0.07) and inter-observer (DSC=0.80; SD=0.12) agreement were excellent.

Manual adaptations of registration errors

Each registration result underwent rigorous quality control to ensure that the infarct projected on the brain template matched the original brain scan in terms of infarct location, size and shape. Manual adaptations were made in 44% of cases (N=527). The most common errors in the registration were: 1) imperfect alignment due to the mass effect caused by the lesion in the acute stage; 2) misalignment of the tentorium cerebelli, in which case an occipital infarct can overlap with the cerebellum in the brain template; 3) misalignment or deformation of periventricular infarcts in patients with enlarged ventricles; and 4) incomplete coverage of cortical areas due to presence of brain atrophy. Manual adaptations were made by an experienced rater (N.A.W) who followed a previously published protocol.[4] An in-house developed brush tool in MevisLab was used to add or remove voxel clusters manually in three-dimensional orientation.[5]
**Definition of stroke subtypes**

Four stroke subtypes were defined: A) Small subcortical infarcts: single supratentorial infarct without cortical involvement, with a lesion volume of $\leq 4.19$ ml (i.e. a sphere of $\leq 2$ cm diameter; following the STRIVE criteria).[6] B) Large subcortical infarcts: supratentorial infarct(s) without cortical involvement, with a lesion volume of $>4.19$ ml. C) Cortical infarcts: supratentorial infarct(s) of any volume with cortical involvement. D) Infratentorial infarcts: any brain stem and/or cerebellar infarct(s).

Patients with multiple infarcts in both supra- and infratentorial regions could be included in both B/C and D. Categories A, B, and C were mutually exclusive. Whether an infarct had cortical or infratentorial involvement was determined using brain masks for the MNI structural atlas (supratentorial cortical regions and cerebellum) [7] and Harvard-Oxford brain atlas (brain stem).[8]
Supplementary Table 1. Examples of cut-off scores for impairment on the MMSE for different ages and education levels

| Age of patient | Years of education | MMSE score | Age- and education-corrected percentile score |
|----------------|--------------------|------------|---------------------------------------------|
| 71 years       | 9 (average)        | 24         | 3rd-5th                                     |
| 60 years       | 9 (average)        | 25         | 3rd-5th                                     |
| 82 years       | 9 (average)        | 23         | 3rd-5th                                     |
| 70 years       | 6 (low)            | 23         | 3rd-5th                                     |
| 72 years       | 16 (high)          | 26         | 3rd-5th                                     |

These real-life examples were taken from the Hallym and Bundang VCI cohorts to illustrate how age and education influenced the cut-off for impairment on the Mini-Mental State Examination. Percentile scores were calculated using Korean normative data.[9] Percentile scores were not corrected for sex and were therefore identical for males and females.
**Supplementary Table 2.** Cognitive profile of the study population and categorization of neuropsychological tests into cognitive domains.

| Cognitive test                                                                 | N available in study sample | Mean (SD; range)          | % impaired |
|-------------------------------------------------------------------------------|-----------------------------|---------------------------|------------|
| Mini-Mental State Examination [9]                                             | 1198                        | 24.01 (5.87; 0-30)        | 35.1       |
| **Attention and executive functions**                                        |                             |                           |            |
| Phonemic fluency (three phonemes, number of words in one minute per phoneme) [10] | 1052                        | 18.8 (11.1; 0-56)         | 18.9       |
| Korean-Trail Making Test - Elderly’s version B (time in seconds, range 0-300) [11] | 884                         | 84.4 (77.8; 11-300)       | 19.5       |
| **Language**                                                                 |                             |                           |            |
| Short Form of the Korean-Boston Naming Test (number correct, range 0-15) [12] | 1170                        | 10.1 (3.6; 0-15)          | 21.6       |
| Semantic fluency, category animals (number of words in one minute) [13]      | 1171                        | 11.8 (5.1; 0-30)          | 23.1       |
| **Processing speed**                                                         |                             |                           |            |
| Korean-Trail Making Test - Elderly’s version A (time in seconds, range 0-300) [11] | 972                         | 38.3 (33.1; 8-300)        | 14.2       |
| Digit Symbol Coding (number correct) [13]                                    | 1050                        | 38.7 (22.2; 0-106)        | 18.6       |
| **Verbal memory†**                                                           |                             |                           |            |
| Seoul Verbal Learning Test - immediate recall (number correct, range 0-36) [14] | 1172                        | 15.1 (6.0; 0-36)          | 29.6       |
| Seoul Verbal Learning Test – delayed recall (number correct, range 0-12) [14] | 1170                        | 3.9 (2.9; 0-12)           | 33.5       |
| Seoul Verbal Learning Test – recognition (number correct, range 0-24) [14]   | 1158                        | 18.2 (4.3; 0-24)          | 33.0       |
| **Visuospatial abilities**                                                   |                             |                           |            |
| Rey Complex Figure Test: Copy (correctly copied elements, range 0-36) [14]    | 1076                        | 26.6 (8.7; 0-36)          | 30.4       |

† Each of the Seoul Verbal Learning Test scores was included as a separate test for the verbal memory domain. Abbreviation: SD, standard deviation.
Supplementary Figure 1. Voxel-based lesion-symptom mapping results of sensitivity analysis and comparison with main results.
This sensitivity analysis was performed on patients without evidence of pre-stroke cognitive impairment (N=704) based on the Informant Questionnaire on Cognitive Decline in the Elderly (IQ-CODE<3.6) to rule out that the observed associations were caused by pre-existing cognitive problems. Results from the main analysis (N=1198; Figure 2 in main text) are also shown as reference. Detail views of the dotted cubes are shown on the right. L=left, R=right.

A-B: Lesion prevalence map of infarcts are shown for the total sample (panel A) and the sensitivity analysis sample (panel B). Only voxels damaged in ≥5 patient are shown. Infarct distributions are comparable, yet brain coverage (i.e. number of voxels included) is lower for the sensitivity analysis (i.e. 49% compared to 71%). This is due to the smaller study sample (N=704 versus N=1198), and the fact that infarcts were smaller in the selected sample of patients with IQCODE<3.6 (median volume in mL (IQR): 2.2 (10.8); N=704) compared to patients with IQCODE≥3.6 (6.1 (36.1); N=198). Consequently, the number of patients with damage in each voxel was lower. To illustrate, the detail view shows that voxels in the left middle cerebral artery territory were damaged in 10-20 patients (color red-dark orange) in the sensitivity analysis, compared to 20-50 patients (light orange-yellow) in the total dataset. This lower brain lesion coverage, combined with the lower prevalence of impairment on the MMSE in the sample of patients with IQCODE<3.6 (i.e. 29% compared to 72% for IQCODE≥3.6), contributed to decreased statistical power for the VLSM analysis.

C-D: Voxel-based lesion-symptom mapping (VLSM) results for impairment on the MMSE, showing the odds ratio for all tested voxels. The sensitivity analysis did not yield any significant voxels (p<0.01), therefore no threshold for statistical significance was applied for visualization. The color indicates the odds ratio (OR) per voxel: orange to red indicates an increased OR for impairment on the MMSE, yellow indicates no association (OR=1), and dark blue indicates a decreased OR. The distribution of relevant voxels in the sensitivity analysis results are similar to the main results, i.e. highest odds ratios (>15) in left frontotemporal regions and the thalamus (see detail view). Of note, certain regions in the frontotemporal lobes that were strongly associated with impairment on the MMSE could not be included in the sensitivity analysis due to insufficient brain coverage, e.g. the anterior part of the left thalamus (see detail view).
Supplementary references

[1] Ricci PE, Burdette JH, Elster AD, et al. A Comparison of Fast Spin-Echo, Fluid-Attenuated Inversion-Recovery, and Diffusion-Weighted MR Imaging in the First 10 Days after Cerebral Infarction. *American journal of neuroradiology*. 1999;20(8):1535-42.

[2] Dice LR. Measures of the Amount of Ecologic Association Between Species. *Ecology* 1945; 26: 297–302.

[3] Crum WR, Camara O, Hill DLG. Generalized Overlap Measures for Evaluation and Validation in Medical Image Analysis. *IEEE Trans Med Imaging* 2006; 25: 1451–1461.

[5] Biesbroek JM, Kuijf HJ, Weaver NA, et al. Brain Infarct Segmentation and Registration on MRI or CT for Lesion-symptom Mapping. *J Vis Exp*. Epub ahead of print September 2019. DOI: 10.3791/59653.

[5] Ritter F, Boskamp T, Homeyer A, et al. Medical image analysis. *IEEE Pulse* 2011; 2: 60–70.

[6] Wardlaw JM, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol* 2013; 12: 822–838.

[7] Collins DL, Holmes CJ, Peters TM, Evans AC. Automatic 3-D model-based neuroanatomical segmentation. *Human brain mapping*. 1995;3(3):190-208.

[8] Makris N, Goldstein JM, Kennedy D, Hodge SM, Caviness VS, Faraone SV, Tsuang MT, Seidman LJ. Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophrenia research*. 2006 Apr 1;83(2-3):155-71.

[9] Kang Y. A Normative Study of the Korean-Mini Mental State Examination(K-MMSE) in the Elderly. *Kor J Psychol*. 2006;25:1-12.

[10] Kang Y, Chin JH, Na DL, Lee J, Park JS. A normative study of the Korean version of Controlled Oral Word Association Test (COWAT) in the elderly. *Korean J Clin Psychol*. 2000;19.2: 385-392.

[11] Yi H, Chin J, Lee B, et al. Development and validation of Korean version of Trail Making Test for elderly persons. *Dement Neurocognitive Disord*. 2007;6.2: 54-66.

[12] Kang Y, Kim H, Psychol DN-KJC, et al. A short form of the Korean-Boston Naming Test (K-BNT) for using in dementia patients. *Korean J Clin Psychol*. 1999;18: 125-138.
[13] Yum T, Park Y, Oh-hashi K, et al. The manual of Korean-Wechsler adult intelligence scale. 1992.

[14] Kang Y, Na D, Hahn S. Seoul neuropsychological screening battery. *Incheon: Human brain research & consulting co.* 2003.