Strategic infarct locations predict post-stroke cognitive impairment: a pooled analysis of individual patient data from twelve acute ischemic stroke cohort studies

Nick A. Weaver, MD[1], Hugo J. Kuijf, PhD[2], Hugo P. Aben, MD[3], Jill Abrigo, MD[4], prof. Hee-Joon Bae, PhD[5], Mélanie Barbay, MD[6], Jonathan G. Best, MD[7], prof. Régis Bordet, PhD[8], Francesca M. Chappell, PhD[9,10], Christopher P.L.H. Chen, MD[11,12], Thibaut Dondaine, PhD[8], Ruben S. van der Giessen, PhD[13], prof. Olivier Godefroy, PhD[6], Bibek Gyanwali, MD[11,12], Olivia K.L. Hamilton, MSc[9,10], Saima Hilal, PhD[11,12,14], Irene M.C. Huenges Wajer, PhD[1,15], prof. Yeonwook Kang, PhD[16], prof. L. Jaap Kappelle, PhD[1], Beom Joon Kim, PhD[5], Sebastian Köhler, PhD[17], Paul L.M. de Kort, PhD[3], prof. Peter J. Koudstaal, PhD[13], Gregory Kuchcinski, MD[8], Bonnie Y.K. Lam, PhD[18,19], prof. Byung-Chul Lee, PhD[20], Keon-Joo Lee, MD[5], Jae-Sung Lim, MD[20], Renaud Lopes, PhD[8], Stephen D.J. Makin, PhD[21], Anne-Marie Mendyk, RN[8], prof. Vincent C.T. Mok, MD[18,19], Mi Sun Oh, MD[20], prof. Robert J. van Oostenbrugge, PhD[22], Martine Roussel, PhD[6], Lin Shi, PhD[4,23], Julie Staals, PhD[22], Maria del C. Valdés-Hernández, PhD[9,10], Narayanaswamy Venketasubramanian, MBBS[24], prof. Frans R.J. Verhey, PhD[17], prof. Joanna M. Wardlaw, PhD[9,10], prof. David J. Werring, PhD[7], Xu Xin, PhD[11,12], prof. Kyung-Ho Yu, PhD[20], Martine J.E. van Zandvoort, PhD[1,15], Lei Zhao, PhD[23], J. Matthijs Biesbroek, PhD[1], prof. Geert Jan Biessels, PhD[1]

[1]Department of Neurology and Neurosurgery, UMC Utrecht Brain Center, Utrecht, the Netherlands

[2]Image Sciences Institute, University Medical Center Utrecht, Utrecht, the Netherlands

[3]Department of Neurology, Elisabeth Tweesteden Hospital, Tilburg, the Netherlands

[4]Department of Imaging and Interventional Radiology, The Chinese University of Hong Kong, Hong Kong SAR, China

[5]Department of Neurology, Seoul National University Bundang Hospital, Seoul National University College of Medicine, Seongnam, Republic of Korea

[6]Department of Neurology, Amiens University hospital, Laboratory of Functional Neurosciences (UR UPJV 4559), Jules Verne Picardy University, 80054 Amiens Cedex, France

[7]Stroke Research Centre, Department of Brain Repair and Rehabilitation, UCL Queen Square Institute of Neurology, Russell Square House, 10 - 12 Russell Square, London WC1B 5EH, United Kingdom.

[8]Université Lille, Inserm, CHU Lille, U1172 - LilNCog - Lille Neuroscience & Cognition, F-59000 Lille, France

[9]Neuroimaging Sciences, Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, United Kingdom

[10]UK Dementia Research Institute at the University of Edinburgh, Edinburgh, United Kingdom

[11]Department of Pharmacology, National University of Singapore, Singapore, Singapore

[12]Memory, Aging and Cognition Center, National University Health System, Singapore, Singapore

[13]Department of Neurology, Erasmus Medical Center, Rotterdam, the Netherlands

[14]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore, Singapore

[15]Experimental Psychology, Helmholtz Institute, Utrecht University, the Netherlands

[16]Department of Psychology, Hallym University, Chuncheon, Republic of Korea

[17]Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University, Maastricht, the Netherlands

[18]Division of Neurology, Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong SAR, China

[19]Therese Pei Fong Chow Research Centre for Prevention of Dementia, Margaret Kam Ling Cheung Research Centre for Management of Parkinsonism, Gerald Choa Neuroscience Centre, The Chinese University of Hong Kong, Hong Kong SAR, China

[20]Department of Neurology, Hallym University Sacred Hospital, Hallym Neurological Institute, Hallym University College of Medicine, Anyang, Republic of Korea

[21]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, United Kingdom

[22]Department of Neurology, Maastricht University Medical Center, Maastricht, the Netherlands

[23]BrainNow Research Institute, Shenzhen, Guangdong Province, China

[24]Raffles Neuroscience Centre, Raffles Hospital, Singapore, Singapore

Corresponding author: G.J. Biessels, Department of Neurology G03.129, University Medical Center Utrecht, Heidelberglaan 100, 3508 GA Utrecht, the Netherlands. E: g.j.biessels@umcutrecht.nl; T: +31887556866

Abstract

Background: Post-stroke cognitive impairment (PSCI) occurs in approximately half of ischemic stroke survivors. Infarct location is a potential determinant of PSCI, but a comprehensive map of strategic infarct locations is lacking. In this large-scale multicenter lesion-symptom mapping study, we aimed to identify infarct locations most strongly predictive of PSCI, and use this information to develop a prediction model.

Methods: We harmonized individual patient data from twelve cohorts through the Meta-VCI-Map consortium. Patients with acute symptomatic infarcts on CT/MRI and cognitive assessment <1 year post-stroke were eligible. PSCI was defined as impairment in ≥1 cognitive domains on neuropsychological assessment or impairment on the Montreal Cognitive Assessment. Voxel-based lesion-symptom mapping (VLSM) was used to calculate voxel-wise odds ratios for PSCI. For the prediction model, a "location impact score" on a five-point scale was derived from the VLSM results. Combined internal-external validation was performed using leave-one-cohort-out cross-validation for all twelve cohorts.

Findings: In our combined sample of 2950 patients (age 67±12 years, 39% female), 44% had PSCI. We achieved almost complete lesion coverage of the brain in our analyses (87%). Infarcts in the left frontotemporal lobes, left thalamus, and right parietal lobe were strongly associated with PSCI (False Discovery Rate corrected $q<0{\cdot}01$; voxel-wise odds ratios >20). These strategic regions were mapped onto a three-dimensional brain template to visualize PSCI risk per brain region. The location impact score showed good correspondence between predicted and observed risk across cohorts after adjusting for cohort-specific PSCI occurrence.

Interpretation: This study provides the first comprehensive map of strategic infarct locations associated with risk of PSCI. A location impact score was derived from this map that robustly predicted PSCI across cohorts and can be applied by clinicians to identify individual patients at risk of PSCI.

Funding: ZonMw-Vici 918.16.616.

Research in Context

Evidence before this study

We systematically searched PubMed for original studies (no case reports) and systematic reviews on infarct characteristics on brain imaging as predictors of post-stroke cognitive impairment (PSCI), published between January 1, 2000, and November 6, 2020, without language restriction, using the terms ("cognitive impairment" or "dement$"), ("ischemic stroke" or "ischaemic stroke" or "infarct"), and ("CT" or "MRI"). We searched for studies with cognitive impairment or dementia as outcome, rather than studies on neuroanatomical correlates of specific cognitive functions. We identified 17 original studies, and one systematic review from 2009. Cortical involvement, presence of multiple acute infarcts, total infarct volume, and left cerebral hemispheric location (versus right) were consistently associated with PSCI. Seven of these studies assessed associations between specific infarct locations and PSCI. Infarct locations significantly associated with PSCI included the left thalamus and basal ganglia, but results were not consistent. This is likely attributable to different anatomical categorizations, mostly with low spatial resolution. Moreover, the two previous studies that evaluated lesion location at voxel level had limited brain coverage (i.e. less than 20% of total brain volume), leaving a major proportion of the brain unexplored.

Added value of this study

In this pooled analysis of patient data from 2950 patients from 12 cohorts, we were able to include 87% of the entire brain, thus essentially covering all possible infarct locations. The large sample size enabled us to perform stratified analyses for different stroke subtypes, which is essential considering associated differences in lesion size and distribution. We also pinpointed unique patterns for different cognitive domains. The large-scale multicenter design enhances generalizability of the results, as supported by our leave-one-cohort-out cross-validations. These results formed the basis for a prediction score that allows for a robust and individualized PSCI risk assessment in the clinical setting, solely based on a patient's infarct location on brain CT or MRI.

Implications of all the available evidence

Building upon the historical notion that infarct location determines cognitive impairment, we now firmly established which locations are most strongly predictive of PSCI across almost the entire brain. We provide an easy-to-use visual rating scale that can be applied by clinicians to identify patients at risk. This could facilitate cognitive prognostication, help to identify patients that should undergo cognitive assessment, and allow for timely intervention with rehabilitative treatment strategies. The visual rating scale is freely available on our website (https://metavcimap.org/features/software-tools/location-impact-score/).

Introduction

Post-stroke cognitive impairment (PSCI) is a common consequence of ischemic stroke, and a leading cause of long-term disability and reduced quality of life.[1] PSCI occurs in approximately half of stroke survivors within the first year,[2,3] yet it is difficult to provide patients with an individualized cognitive prognosis. Stroke characteristics, including lesion burden on brain imaging, have been shown to predict PSCI to a greater extent than demographic and vascular risk factors.[2,4] However, more specific imaging predictors that could identify patients at risk of PSCI are lacking.[4,5]

Infarct location is widely recognized as a determinant of PSCI. For example, the left angular gyrus and thalamus are considered to be strategic infarct locations for PSCI, but this knowledge was largely based on small case series.[6,7] In the past two decades, lesion-symptom mapping techniques have emerged that can study the relationship between lesion location and cognition in large datasets, with high spatial resolution.[8] These techniques can elucidate the neuroanatomy of cognitive functions, but can also be used for predictive modeling.[9] Initial studies applying this technique have confirmed the left angular gyrus and thalamus as strategic locations for PSCI, and pinpointed additional locations.[10,11] However, despite sample sizes of several hundred patients, these studies could cover less than 20% of the brain in their analyses.[12] To overcome this limitation, the Meta-VCI-Map consortium was initiated, to improve lesion coverage by pooling data from multiple cohorts . A pilot study showed successful harmonization of clinical and imaging data from six cohorts (N=878), achieving 46% brain coverage, indicating that two to three thousand patients would be required for complete coverage.[12]

In this Meta-VCI-Map consortium study, we performed a large-scale multicenter lesion-symptom mapping study in 2950 patients with acute ischemic stroke from 12 cohorts, aiming to identify infarct locations most strongly predictive of PSCI and create a comprehensive brain map of strategic infarct locations. Furthermore, we derived a 'location impact score' from this map to develop a prediction model to identify patients at risk of PSCI.

Methods

Study design and participants

Eligible cohorts were identified (per January 2019) through the Meta-VCI-Map consortium (www.metavcimap.org)[12] with the following criteria: (1) cohort of patients with acute ischemic stroke; (2) brain CT or MRI (T1/T2/FLAIR/DWI) showing the symptomatic infarct(s); (3) available infarct segmentations; and (4) cognitive assessment ≤1 year post-stroke, consisting of either multidomain neuropsychological assessment or Montreal Cognitive Assessment (MoCA). Patients with a history of stroke could be included. Twelve cohorts from France (N=2), Hong Kong (N=1), the Republic of Korea

(N=2), the Netherlands (N=4), Singapore (N=1), and the United Kingdom (N=2) participated (Table 1). Most studies excluded patients with severe dysphasia based on clinical examination prior to inclusion, because most cognitive assessment protocols were incompatible with dysphasia (appendix p.20). For the combined analysis in the present study, we did not further select on language performance. Central data processing and analysis were performed at the University Medical Center Utrecht.

Cognitive data harmonization

For the nine cohorts with multidomain neuropsychological assessment, PSCI was defined as "cognitive impairment in one or more cognitive domains", in accordance with the VASCOG criteria for Vascular Cognitive Disorders.[13] Cognitive tests from each cohort were assigned to six domains: attention and executive functioning; information processing speed; language; verbal memory; visuospatial perception/construction; and visuospatial memory. Only tests with (local) norm-referenced data were used (overview per cohort in appendix, p.22-24). Assignment to domains was determined by a neuropsychology working group (members: NAW, IMCHW, OG, OKLH, AW, BYKL, JSL, TD, and XX). For each test, performance <5th percentile was defined as impaired. Performance on a cognitive domain was impaired if >50% of available tests on that domain were impaired, determined on a per-subject basis. Data on a minimum of three domains was needed to rule out PSCI.

For the three cohorts in which only the MoCA was available, PSCI was defined as a performance $<5^{th}$ percentile on the total MoCA score, using age- and education-corrected normative data (appendix p.25).

Lesion data harmonization

Lesion segmentations of acute symptomatic infarcts were provided by participating centers. Both CT and MRI scans were used for infarct segmentation, depending on their availability. We adhered to previously published protocols regarding suitable time windows for delineating acute infarcts.[14] Pre-existing infarcts and other vascular lesions were not analyzed. Infarct segmentation methods are listed in the appendix (p.26). Lesion maps were registered to the 1-mm MNI-152 brain template for spatial normalization.[15] For 11 cohorts, lesion registration was performed with the RegLSM image processing tool (available at www.metavcimap.org).[14] RegLSM provides custom-fit settings for CT and MRI (T1/FLAIR/DWI) and can reliably process scans from different vendors and field strengths to a uniform output, as was demonstrated in the Meta-VCI-Map pilot study,[12] allowing lesion data from different sources to be combined into one dataset. The GRECogVASC study used the pyramidal block-matching Diffeomorphic Demons algorithm in the MedINRIA software package for lesion registration,[16] which provides comparable output as RegLSM.

For nine cohorts, lesion maps were already registered to MNI-152 space as part of previous projects. For the remaining three cohorts (CASPER, CROMIS-2 and STROKDEM), registrations were generated in the current project, with visual inspection of the registration results by an experienced rater (NAW). Manual adaptations were made in case of minor displacements (appendix p.2). Rigorous quality control of fully processed lesion data was performed for all 12 cohorts (appendix p.2).

Voxel-based lesion-symptom mapping

Voxel-based lesion-symptom mapping (VLSM) was performed to relate infarct location to PSCI occurrence. Voxels damaged in <5 patients were excluded from all analyses. Voxel-wise odds ratios (OR) for PSCI were calculated using the Fisher's exact test in Python (SciPy 1.4.1). False discovery rate correction (FDR; threshold $q<0{\cdot}01$) was applied to correct for multiple comparisons. In addition, voxel-wise absolute risks (i.e. percentage of patients with PSCI for each particular voxel) were calculated.

Several supplementary analyses were performed using the same VLSM approach: analyses with each of the six cognitive domains as outcome, and stratified analyses for three stroke subtypes (small subcortical infarcts, cortical or large subcortical infarcts, and infratentorial infarcts (definitions in appendix, p.2)). Sensitivity analyses were performed on patients with cognitive assessment ≥2 weeks post-stroke (i.e. to limit the effect of factors in the acute stage (e.g. delirium) on cognition), and patients with multidomain neuropsychological assessment (i.e. excluding patients with only MoCA).

Location impact score and prediction models

A "location impact score" was calculated for each patient based on the mean coefficient (i.e. ln(OR)) of voxels from the VLSM results of the patient's infarct (appendix, p.4). For ease of use and visual rating, we re-coded this continuous score into a five-point score based on quintiles. The association between the continuous and five-point variants of the location impact score and PSCI was assessed with logistic regression analysis in univariate models, and subsequently in multivariate models together with known clinical PSCI predictors (based on literature[4]): age, sex, education level (categories in appendix, p.28), time interval between stroke onset and cognitive assessment, history of stroke, and total infarct volume in MNI-152 space (i.e. corrected for brain volume).

We performed a combined internal-external validation of the location impact score (continuous and five-point) using leave-one-cohort-out cross-validation for all twelve cohorts to assess generalizability, robustness, and cohort-specific effects.[17] Furthermore, we explored the added value of the location impact score on top of the known clinical PSCI predictors in multivariate logistic regression models. A base model was built using the clinical predictors only, followed by addition of total infarct volume, and finally the location impact score.

Goodness-of-fit was taken as the main feature of interest, which we assessed by plotting deciles of predicted PSCI risks versus observed occurrence in calibration plots. The c-statistic was used to assess model discrimination, though it is known to underestimate the prognostic value of individual biomarkers.[18] Prediction models were built using R version 3.6.3 (packages: "glmnet" and "rms").

Visual rating of the location impact score

To enable visual assessment of the five-point location impact score, a visual rating scale was created by grouping voxels from the VLSM results into five color categories based on continuous location impact score quintiles. Three clinicians (two neurologists [GJB/JMB]) and a neurology resident [MC]) tested reliability of visual rating of scans against the computer-generated five-point scale. Raters were provided with 12 brain template slices showing the scale (appendix, p.5). Raters evaluated a random subset of 36 MRI scans, selected from 3 cohorts, by estimating the average score of the region affected by the infarct (DWI: N=28; FLAIR: N=8), blinded to clinical data. Agreement between the visual rating and the actual score and inter and intra-rater agreement were assessed with Cohen's weighted Kappa (range: 0-1, where 1 is perfect agreement).[19] Disagreements were weighted according to their squared distance from perfect agreement.

Ethics statement

For all cohorts, ethical and institutional approval was obtained as required by local regulations to allow data acquisition, including informed consent, and data sharing. All data processed at the coordinating center for the pooled analysis was anonymized.

Role of the funding source

The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report, and had no involvement in the decision to submit the paper for publication.

Results

Sample characteristics

We included 2950 patients from 12 cohorts (Figure 1). Participants had heterogeneous clinical characteristics, reflecting the different study designs of the cohorts (Table 1; appendix, p.29). In the total study sample, mean age was 67 years (SD 12) and 1157 patients (39%) were female. Approximately half of the study population was Korean (47%) and one-third (36%) was European Caucasian. Cognitive assessment was generally performed 2-6 months post-stroke (median 105 days, IQR 74-170). MRI DWI

(66%) and T2/FLAIR (16%) were most commonly used for infarct segmentation. Imaging was mostly performed in the first two weeks post-stroke (median 4 days, IQR 1-9).

Cognitive profile

Nine cohorts used multidomain neuropsychological assessment (N=2343, 79%), and three cohorts only the MoCA (N=607, 21%). PSCI occurrence was 44% in the total sample (appendix p.6, panel 3A), with higher occurrence in cohorts with multidomain neuropsychological assessment (range 18-62%) than those with MoCA (3-20%). Cohorts that by design excluded patients with severe stroke (STROKDEM and Mild Stroke Study 2) had a lower PSCI occurrence. Cognitive domains were affected in variable frequencies (range in total sample 12-27%; appendix p.6, panel 3B).

Brain lesion coverage

In the combined sample, infarct distribution was symmetrical and brain lesion coverage was high: 87% of brain voxels were damaged in ≥5 subjects and could therefore be included in the analyses (Figure 2A). Only the most distal part of the anterior cerebral artery territories, parts of the midbrain, and temporal poles could not be included in the analyses due to infrequent involvement. Lesion prevalence maps for the individual cohorts are provided in the appendix (p.7). In the individual cohorts, infarct patterns reflected cohort-specific inclusion criteria.

Voxel-based lesion-symptom mapping results

Figure 2B-C shows ORs for PSCI occurrence in the total sample (N=2950). Particularly,infarcts in the left frontotemporal lobes, right parietal lobe, and left thalamus were significantly associated with PSCI after correction for multiple testing (FDR q<0·01), with high voxel-wise ORs (>20). Over 80% of patients with infarcts in these regions had PSCI (appendix, p.8). By contrast, infarcts in the right occipital lobe, brain stem and cerebellum conveyed a lower risk of PSCI (OR<1; not statistically significant) relative to infarcts elsewhere. Sensitivity analyses limited to patients with cognitive assessment ≥2 weeks post-stroke (N=2554) and patients with multidomain neuropsychological assessment (N=2343) showed comparable results (appendix, p.9).

Figure 2D-F shows the VLSM results for the different stroke subtypes (prevalence maps in appendix, p.10). In patients with cortical and large subcortical infarcts (N=1488), locations associated with PSCI were the same as in the total sample. For patients with small subcortical infarcts (N=901) the highest ORs for PSCI (>20) were found in the left thalamus. For patients with infratentorial infarcts, no infratentorial locations were associated with elevated PSCI risk (N=708). Based on these results, domain-specific analyses were only performed in the subgroup with cortical or large subcortical infarcts. This revealed differential infarct

patterns across both hemispheres for different domain-specific impairments (Figure 3). For example, ORs for impairment of processing speed showed limited variation across the brain, without clear associations with particular infarct locations. By contrast, high ORs for language and verbal memory impairment were strongly lateralized to the left hemisphere, whereas impairment of visuospatial functions was associated with right parietal infarcts.

Prediction models

Both the continuous and five-point variants of the location impact score were significantly associated with PSCI in the total dataset (Table 2), also independent from other predictors (appendix, p.31). Leave-one-cohort-out cross-validation of the continuous (appendix, p.11-14) and five-point location impact score (Figure 4) revealed good correspondence between predicted and observed risk in four well-powered cohorts (Bundang VCI, CU-STRIDE, Hallym VCI, PROCRAS) after adjusting the intercept for cohort-specific PSCI occurrence. C-statistics for these four cohorts ranged from 0·55 (95%CI: 0·51-0·60) to 0·62 (0·52-0·71) and were similar for the continuous and five-point scores. Five cohorts (CASPER, COAST, CROMIS-2, STROKDEM, USCOG) showed a similar trend, but they provided less stable estimates most likely related to limited sample size and/or low PSCI occurrence. One well-powered cohort showed poor model calibration and discrimination (GRECogVASC). Two cohorts (CODECS and MSS-2) could not be used as validation sample because of insufficient sample size or low PSCI occurrence (appendix, p.11-14). Cross-validation of the multiple logistic regression model containing the known clinical PSCI predictors, but without the location impact score, had poor external performance, which was slightly improved by addition of total infarct volume (appendix, p.15-18). Addition of the location impact score provided the best correspondence between predicted and observed risk, across a much wider range of predicted probabilities. Of note, the location impact score by itself showed similar performance as the combined model (appendix, p.11-18), therefore the univariate model with only the location impact score was selected as final model (Table 2).

Visual rating of the location impact score

A color map was created of the five-point location impact score for visual assessment (appendix, p.19; https://metavcimap.org/features/software-tools/location-impact-score/). Correspondence between visual ratings and the actual location impact score was excellent for all three raters (Cohen's weighted Kappa: range 0·88-0·92). Inter-rater agreement (range 0·85-0·87) and intra-rater agreement (0·95) were also excellent. Visual rating typically took 1-2 minutes per scan.

Discussion

This large-scale multicenter study provides the first comprehensive map of strategic infarct locations predicting PSCI. With data from 2950 patients we achieved almost complete coverage of the brain, and showed that infarcts in the left frontotemporal regions, the right parietal lobe and the left thalamus are most strongly predictive of PSCI. From the three-dimensional strategic infarct map we derived a location impact score and validated a single-variable prediction model for PSCI, which showed robust performance across cohorts despite heterogeneity in study populations, supporting generalizability of the findings.

The relationship between brain lesion location and cognition has been widely studied, dating back to early case series from the 19th century that described individuals with a specific deficit who had a lesion at a particular location.[7] Although these historical observations are clearly of value, their generalizability is uncertain due to potential selection and publication bias. In the past two decades, over one-hundred studies applied modern lesion-symptom mapping techniques to analyze associations between infarct locations and cognition at a group level.[20] The vast majority of these studies aimed to elucidate the neuroanatomical correlates of specific cognitive functions. By contrast, only few studies assessed the predictive value of infarct location for PSCI as clinical outcome, which requires a different analytical approach. A large meta-analysis found that left hemisphere lateralization was a significant predictor of PSCI, but this provided little anatomical precision.[4] Only one study thus far combined VLSM with predictive modelling and found that infarcts in left frontotemporal regions, hippocampus, and thalamus were predictive of PSCI[9], in line with our findings. However, this study lacked brain coverage due to the modest sample size (N=267), and results were not externally validated. We overcame these limitations in our pooled dataset achieving 87% brain coverage, allowing us to evaluate and externally validate the predictive value of almost every infarct location across the brain.

Previous lesion-symptom mapping studies mostly used cognitive functioning, often on a single domain, as a continuous outcome measure. Our PSCI definition, in concordance with internationally established diagnostic criteria,[13] was chosen to be compatible with PSCI diagnosis in daily clinical care, which is by nature a dichotomous outcome (i.e. PSCI is present or absent). Availability of multidomain cognitive data also allowed us to also zoom in on infarct locations associated with specific domains. We established that strategic infarct locations that predict PSCI were representative for various cognitive domains (Figure 5), showing that these strategic infarct locations indeed predict PSCI in a broad sense. Domain-specific analyses also revealed distinct patterns compatible with prior knowledge (e.g.[11,12,21]), which can be used for domain-specific prediction.

Stroke etiology (e.g. small versus large vessel disease) and lesion size and volume (e.g. lacunar versus non-lacunar) are also related to PSCI,[4] yet are inextricably linked to lesion location. For supratentorial infarcts, small subcortical infarcts inherently had a different infarct distribution than cortical and large subcortical

infarcts. The left thalamus was predictive of PSCI in both stroke subtypes, while other subcortical regions (e.g. left basal ganglia) were only predictive of PSCI as part of a larger infarct.

When interpreting our findings, it is important to consider that infarct locations predictive of PSCI do not necessarily imply causality. When lesion-symptom mapping is used to determine neuroanatomical correlates of cognitive deficits, it is important to take confounding factors such as the interdependency between infarct locations (i.e. infarct patterns are not random but follow the vascular tree), clinical history (e.g. pre-stroke cognition, depression), and other brain lesions (e.g. white matter hyperintensities, previous infarcts) into account.[8] Meanwhile, for prediction, adding variables to a model can improve predictive ability, but confounding is by definition not an issue. We did explore if a combined model with other known predictors of PSCI would perform better than the univariate model with the location impact score only, but found that the univariate model was equally robust in cross-validations. The final model, consisting of the location impact score only, allows for easy clinical implementation due to its simplicity. It is likely that with future developments of harmonization and quantification of imaging markers, such as white matter hyperintensities[4,10,22] previous infarcts,[4] and brain atrophy,[4,10] more elaborate PSCI prediction models can be developed, although clinical implementation may be more challenging.

The location impact score is the first-ever prediction tool for PSCI based solely on clinically available brain images. While awaiting further testing in prospective studies, this score appears promising for direct application in clinical practice for cognitive prognostication. Visual ratings were fast and reliable, even without prior training. We envision that this tool can be applied upon hospital admission for an early assessment of PSCI risk, by using either a printed (appendix, p.5) or digital version (https://metavcimap.org/features/software-tools/location-impact-score/) and comparing this to the patient's brain scan. This could help with early identification of patients at risk of PSCI, prioritize patients for cognitive assessment where resources are restricted, provide better information to patients and caregivers during planning of aftercare, and allow for timely intervention with rehabilitative treatment strategies. The location impact score could also be used to identify high-risk patients for intervention trials. Of note, in patients who according to the location impact score are at relatively low risk, PSCI occurrence was still up to 30% (Table 2). This implies that the score should not be used to rule out PSCI, and that even in locations with relatively low predicted PSCI risk (e.g. the cerebellum) PSCI is not uncommon. On the other hand, the location impact score might also underestimate PSCI risk in some patients: the predicted PSCI occurrence for score 5 is 66% (Table 2), while >80% of patients with infarcts in certain locations actually had PSCI (appendix, p.8).

We included multicenter data that enabled a much larger sample size and geographical coverage than individual studies to date. This inherently resulted in heterogeneity. Lesion data was based on different modalities and sequences, acquired at different post-stroke time intervals. Our processing pipeline was

adapted to deal with this heterogeneity, although timing and modality of imaging might affect lesion features such as size and resolution. Neuropsychological test batteries also differed between cohorts, but our PSCI definition did allow us to cover a broad range of cognitive profiles and deficits. To account for cohort-specific effects in our analyses and determine the external validity and generalizability, we performed extensive cross-validations that showed robust prediction of PSCI across heterogeneous cohorts. Notably, model fit was mainly influenced by differences in PSCI occurrence, related to differences between cohorts in the sensitivity of cognitive measures and preselection of patients (e.g. exclusion of dementia).

Some potential limitations must also be noted. First, cohort studies on PSCI tend to be selective. For example, to be eligible for cognitive testing, (severe) dysphasia was generally an exclusion criterion. Hence, patients with language impairment are relatively underrepresented in our sample, and the association with language-related brain regions might be attenuated. Despite this selectivity, our results still indicated a major contribution of left hemispheric regions that are associated with language ability, therefore exclusion of these patients is unlikely to have caused missing of key brain regions. Second, post-stroke cognitive trajectories differ between patients and may change over time. Therefore variation in timing of assessment might influence our results.[5] Third, information on pre-stroke cognitive status was not available for a substantial number of cohorts, therefore we could not take this into account in our analyses. Finally, our pooled sample only includes Caucasian and Asian patients, thus generalizability to other ethnicities remains undetermined.

In summary, this study provides the first comprehensive map of strategic infarct locations associated with a high risk of PSCI. A location impact score was derived from this map that robustly predicted PSCI across cohorts and can be applied by clinicians to identify individual stroke patients at risk of PSCI.

Data availability

The Meta-VCI-Map consortium is dedicated to data sharing. The combined dataset of the present study, including deidentified participant data and lesion maps in MNI space, can be used for additional analyses, upon reasonable request. This requires approval of the project group, consisting of representatives of all participating cohorts. Participating cohorts have shared their data with the consortium under the agreement that the source data remains property of the group that supplied the data. Source data are stored at the participating centers and can be requested there.

Acknowledgements

The authors thank Maria A. de Winter and Lieza G. Exalto for their input on the prediction models, Adrian Wong (AW) for his advice on the neuropsychological data harmonization, and Mirthe Coenen (MC) for testing the visual rating scale.

Funding

The Meta VCI Map consortium is supported by Vici Grant 918.16.616 from ZonMw, The Netherlands, Organisation for Health Research and Development, to GJB. Analyses for this study are supported by a Rudolf Magnus Young Talent Fellowship from the University Medical Center Utrecht Brain Center to JMB.

Additional funding support of individual cohorts: CASPER: supported by Maastricht University, Health Foundation Limburg, and the Adriana van Rinsum-Ponsen Stichting. CROMIS-2: funded by The Stroke Association and the British Heart Foundation (TSA BHF 2009/01). CU-STRIDE: financially supported by the Health and Health Services Research Fund (0708041) of the Food and Health Bureau of the Government of the HKSAR, the Lui Che Woo Institute of Innovative Medicine, and Therese Pei Fong Chow Research Centre for Prevention of Dementia (in memory of Donald H. K. Chow). GRECogVASC: funded by Amiens University Hospital and by a grant from the French Ministry of Health (DGOS R1/2013/144). MSS2: funded by the Wellcome Trust (WT088134/Z/09/A, SM) and the Row Fogo Charitable Trust. OKLH is funded by the College of Medicine and Veterinary Medicine at the University of Edinburgh, and is supported by the Wellcome Trust through the Translational Neuroscience PhD program at the University of Edinburgh. PROCRAS: funding was obtained through ZonMW as part of the "TopZorg" project in 2015 (grant # 842003011).

Contributors

NAW, HJK, JMB, and GJB were responsible for study concept and design. NAW, JMB, and GJB drafted the manuscript. NAW and HJK were responsible for central image processing and lesion data harmonization. NAW and IMCHW were responsible for neuropsychological data harmonization, with critical input from OG, OKLH, BYKL, AW, JSL, TD, and XX. NAW and JMB were responsible for data analysis, with critical input from GJB, HJK, OG, JMW, and FMC.

NAW and JMB have accessed and verified the underlying data.

All authors revised the manuscript critically for important intellectual content. All authors gave final approval for submission.

Cohort-specific contributions: HJB, BJK and KJL were responsible for the Bundang VCI cohort. HJB: data collection and data interpretation; BJK and KJL: data collection. SK, FRJV, JS and RvO were responsible for the CASPER cohort study design and data collection. CPLHC, BG, SH, NV and XX were responsible for the COAST cohort. CPLHC and NV: study design, data collection, data analysis, data interpretation; BG, SH and XX: data collection and interpretation. RvdG and PJK were responsible for the CODECS cohort study design and data collection. JGB and DJW were responsible for the CROMIS-2 cohort. DJW: study design and data collection. JGB: data collection and analysis. BYKL, JA, LS and LZ were responsible

for the CU-STRIDE cohort. VCTM: study concept and design. BYKL, JA, LS and LZ: data collection, image processing and data analysis. VCTM, BYKL, JA, LS and LZ: critical evaluation and interpretation of the results. OG, MR and MB were responsible for the GRECogVASC cohort. MR had a major role in the acquisition of data. OG and MR contributed to the design and conceptualization of the study. OG, MR, and MB analyzed and interpreted the data. JSL, KHY, BCL, MSO, and YK were responsible for the Hallym VCI cohort. JSL: data collection and data interpretation; KHY, BCL, MSO, and YK: data collection. JMW, SM, FC, MvH, and OKLH were responsible for the MSS2 cohort. JMW, SM, FC, MvH were responsible for cohort recruitment, analysis, funding and overall organization. OKLH assisted with data preparation, study concept, design, and harmonization. PK and HPA were responsible for the PROCRAS cohort study design and data collection. RB, TD, RL, GK, and AMM were responsible for the STROKDEM cohort study design, data collection and analysis. LJK, MJEvZ and JMB were responsible for the USCOG cohort. LJK, MJEvZ: study design and data collection. JMB: data collection and analysis.

Conflicts of interest

HJB reports grants from BMS Korea, grants and personal fees from Shire Korea, grants from Chong Gun Dang Pharmaceutical Corp., grants from Dong-A Pharmaceutical, grants from AstraZeneca Korea, grants from Bayer, grants from Daiichi-Sankyo, grants from Yuhan Corporation, grants from Jeil Pharmaceutical Co, grants from Korean Drug Co., Ltd, grants from Shinpoong Pharm. Co. Ltd., grants from Servier Korea, grants from JLK inspection, grants and personal fees from ESAI, personal fees from Amgen Korea, personal fees from Otsuka Korea, and grants from SAMJIN Pharm, outside the submitted work. DJW reports personal fees from Bayer, personal fees from Alnylam, and personal fees from Portola, outside the submitted work. OG reports grants from Amiens University hospital, grants from French Health department, during the conduct of the study; during the last five years he has served on scientific advisory boards and as speaker (Novartis, CSL-Behring, Biogen, Genzyme, Lilly, Bristol-Myers Squibb, Boehringer-Ingelheim, Covidien, Teva and Astra Zeneca), outside the submitted work. MB received funding for travel and meetings from TEVA SANTE SAS, BRISTOL-MYERS SQUIBB, ROCHE SAS, PFIZER SAS, SANOFI ADVENTICE FRANCE, ISIS PERFUSION NORD, and AMGEN SAS. GJB reports grants from ZonMw, The Netherlands, Organisation for Health Research and Development, during the conduct of the study. JMW reports grants from Wellcome Trust, grants from Row Fogo Charitable Trust, grants from Chest Heart Stroke Scotland, grants from Medical Research Council, during the conduct of the study; grants from Fondation Leducq, grants from EU Horizon 2020, grants from British Heart Foundation, grants from Stroke Association, outside the submitted work. HPA reports grants from ZonMW, grant # 842003011, during the conduct of the study.

References

1 Pollock A, St George B, Fenton M, Firkins L. Top ten research priorities relating to life after stroke. *Lancet Neurol* 2012; **11**: 209.

2 Pendlebury ST, Rothwell PM. Incidence and prevalence of dementia associated with transient ischaemic attack and stroke: analysis of the population-based Oxford Vascular Study. *Lancet Neurol* 2019; **18**: 248–58.

3 Barbay M, Taillia H, Nédélec-Ciceri C, *et al.* Prevalence of poststroke neurocognitive disorders using national institute of neurological disorders and stroke-Canadian stroke network, VASCOG criteria (Vascular Behavioral and Cognitive Disorders), and optimized criteria of cognitive deficit. *Stroke* 2018; **49**: 1141–7.

4 Pendlebury ST, Rothwell PM. Prevalence, incidence, and factors associated with pre-stroke and post-stroke dementia: a systematic review and meta-analysis. *Lancet Neurol* 2009; **8**: 1006–18.

5 Mijajlović MD, Pavlović A, Brainin M, *et al.* Post-stroke dementia - a comprehensive review. *BMC Med* 2017; **15**: 1–12.

6 Gorelick PB, Scuteri A, Black SE, *et al.* Vascular contributions to cognitive impairment and dementia: A statement for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2011; **42**: 2672–713.

7 Biesbroek JM, Weaver NA, Biessels GJ. Lesion location and cognitive impact of cerebral small vessel disease. *Clin Sci* 2017; **131**: 715–28.

8 Karnath HO, Sperber C, Rorden C. Mapping human brain lesions and their functional consequences. *Neuroimage* 2018; **165**: 180–9.

9 Munsch F, Sagnier S, Asselineau J, *et al.* Stroke location is an independent predictor of cognitive outcome. *Stroke* 2016; **47**: 66–73.

10 Puy L, Barbay M, Roussel M, *et al.* Neuroimaging determinants of Poststroke cognitive performance: The GRECogVASC Study. *Stroke* 2018; **49**: 2666–73.

11 Zhao L, Biesbroek JM, Shi L, *et al.* Strategic infarct location for post-stroke cognitive impairment: A multivariate lesion-symptom mapping study. *J Cereb Blood Flow Metab* 2017. DOI:10.1177/0271678X17728162.

12 Weaver NA, Zhao L, Biesbroek JM, *et al.* The Meta VCI Map consortium for meta-analyses on strategic lesion locations for vascular cognitive impairment using lesion-symptom mapping: Design and multicenter pilot study. *Alzheimer's Dement Diagnosis, Assess Dis Monit* 2019; **11**: 310–26.

13 Sachdev P, Kalaria R, O'Brien J, *et al.* Diagnostic criteria for vascular cognitive disorders: A VASCOG statement. *Alzheimer Dis Assoc Disord* 2014; **28**: 206–18.

14 Biesbroek JM, Kuijf HJ, Weaver NA, Zhao L, Duering M, Biessels GJ. Brain Infarct Segmentation and Registration on MRI or CT for Lesion-symptom Mapping. *J Vis Exp* 2019; published online Sept. DOI:10.3791/59653.

15 Fonov V, Evans AC, Botteron K, *et al.* Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage* 2011; **54**: 313–27.

16 Arnoux A, Toba MN, Duering M, *et al.* Is VLSM a valid tool for determining the functional anatomy of the

brain? Usefulness of additional Bayesian network analysis. *Neuropsychologia* 2018. DOI:10.1016/j.neuropsychologia.2018.10.003.

17 Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal-external, and external validation. J. Clin. Epidemiol. 2016; **69**: 245–7.

18 Cook NR. Quantifying the added value of new biomarkers: how and how not. *Diagnostic Progn Res* 2018; **2**: 14.

19 Cohen J. A Coefficient of Agreement for Nominal Scales. *Educ Psychol Meas* 1960; **20**: 37–46.

20 Karnath H-O, Rennig J. Investigating structure and function in the healthy human brain: validity of acute versus chronic lesion-symptom mapping. *Brain Struct Funct* 2017; **222**: 2059–70.

21 Biesbroek JM, van Zandvoort MJE, Kuijf HJ, *et al.* The anatomy of visuospatial construction revealed by lesion-symptom mapping. *Neuropsychologia* 2014; **62**. DOI:10.1016/j.neuropsychologia.2014.07.013.

22 Georgakis MK, Duering M, Wardlaw JM, Dichgans M. WMH and long-term outcomes in ischemic stroke: A systematic review and meta-analysis. *Neurology* 2019; **92**: E1298–308.

23 Yu KH, Cho SJ, Oh MS, *et al.* Cognitive impairment evaluated with vascular cognitive impairment harmonization standards in a multicenter prospective stroke cohort in Korea. *Stroke* 2013; **44**: 786–8.

24 Lim J, Kim N, Jang MU, *et al.* Cortical Hubs and Subcortical Cholinergic Pathways as Neural Substrates of Poststroke Dementia. 2014. DOI:10.1161/STROKEAHA.

25 Douven E, Schievink SHJ, Verhey FRJ, *et al.* The Cognition and Affect after Stroke - a Prospective Evaluation of Risks (CASPER) study: rationale and design. *BMC Neurol* 2016; **16**: 65.

26 Dong Y, Venketasubramanian N, Poon-Lap Chan B, *et al.* Brief screening tests during acute admission in patients with mild stroke are predictive of vascular cognitive impairment 3e6 months after stroke. DOI:10.1136/jnnp-2011-302070.

27 Wilson D, Ambler G, Shakeshaft C, *et al.* Cerebral microbleeds and intracranial haemorrhage risk in patients anticoagulated for atrial fibrillation after acute ischaemic stroke or transient ischaemic attack (CROMIS-2): a multicentre observational cohort study. *Lancet Neurol* 2018; **17**: 539–47.

28 Wardlaw JM, Makin SJ, Valdés Hernández MC, *et al.* Blood-brain barrier failure as a core mechanism in cerebral small vessel disease and dementia: evidence from a cohort study. *Alzheimer's Dement* 2017; **13**: 634–43.

29 Aben HP, Reijmer YD, Visser-Meily JM, *et al.* A Role for New Brain Magnetic Resonance Imaging Modalities in Daily Clinical Practice: Protocol of the Prediction of Cognitive Recovery After Stroke (PROCRAS) Study. *JMIR Res Protoc* 2018; **7**: e127.

30 Bournonville C, Hénon H, Dondaine T, *et al.* Identification of a specific functional network altered in poststroke cognitive impairment. *Neurology* 2018; **90**: e1879–88.

Figure captions and legends

Figure 1. Flow chart of patient selection

Patient selection is shown for each cohort separately. Cohorts could join at any given step of data processing. The dark blue boxes indicate at which processing stage the imaging data was provided. The starting number for each cohort indicates how many patients fulfilled the inclusion criteria for the present study. [a] Minimal required clinical data for inclusion were: age, sex, level of education, and either multidomain neuropsychological assessment (with data on 3 or more cognitive domains) or Montreal Cognitive Assessment (MoCA) performed within the first year (maximum: 456 days). Other variables were not essential for inclusion, and missing data has been noted in Table 1. [b] Data previously used for the Meta VCI Map pilot study; selection steps for these datasets have been described previously.[12]

Figure 2. Lesion prevalence maps and voxel-based lesion-symptom mapping results for PSCI

A: Lesion prevalence map of infarcts in the total sample (N=2950), only including voxels damaged in ≥5 subjects. The color indicates how many patients had an infarct in a particular voxel. B-C: Voxel-wise odds ratios (ORs) for post-stroke cognitive impairment (PSCI) occurrence are shown for the total sample (N=2950), calculated using the Fisher's exact test. The color indicates the OR per voxel: dark green to blue indicates that presence of an infarct in that voxel is associated with an increased OR for cognitive impairment compared to absence of an infarct in that voxel, lime green indicates no association (OR=1), and yellow indicates a decreased OR. Panel B shows the ORs for all tested voxels (i.e. damaged in N≥5; no threshold for statistical significance). Panel C only shows voxels with q<0·01 after False Discovery Rate correction. D-F: Voxel-wise odds ratios (ORs) for PSCI occurrence are shown for subgroups of patients with small subcortical infarcts (panel D; PSCI occurrence: 37%), large subcortical or cortical infarcts (panel E; PSCI occurrence: 49%); and infratentorial infarcts (panel F; PSCI occurrence: 43%). Small subcortical infarcts were defined as single supratentorial infarcts without cortical involvement, with lesion volume of ≤4.19 ml (i.e. a sphere of ≤2 cm diameter). Other supratentorial infarcts were categorized as large subcortical (>4.19 ml) or cortical infarcts (any volume). Infratentorial infarcts included brain stem and cerebellar infarcts (any volume). If a patient had both supra- and infratentorial infarcts (in 147 cases), the entire infarcted area was included; hence, some supratentorial regions were included in the infratentorial subgroup analysis (panel F), and vice versa (panel E). The "plasma" and "viridis" scale from the viridis color palette was used for visualization of lesion prevalence and ORs respectively (based on https://cran.r-project.org/web/packages/viridis/index.html). Coordinates of the MNI-152 template (Z; axial orientation) are indicated at the top of the figure. FDR = false discovery rate, L = left, R = right, PSCI = post-stroke cognitive impairment.

Figure 3. Voxel-based lesion-symptom mapping results for domain-specific impairment

This supplementary analysis was performed with domain-specific impairment in six cognitive domains as outcome, to determine to what extent the strategic infarct locations identified for PSCI (i.e. as overarching construct) would represent individual cognitive domains. Voxel-wise odds ratios (ORs) for domain-specific impairment are shown, calculated using the Fisher's exact test. Panel A shows the ORs for all tested voxels (i.e. damaged in N≥5; no threshold for statistical significance). The color indicates the OR per voxel: dark green to blue indicates that presence of an infarct in that voxel is associated with an increased OR for cognitive impairment compared to absence of an infarct in that voxel, lime green indicates no association (OR=1), and yellow indicates a decreased OR. Panel B only shows odds ratios for voxels with q<0·01 after False Discovery Rate correction. Infarct locations associated with PSCI (top row of each panel) corresponded with infarct locations associated with individual cognitive domains, for example the left frontotemporal lobes and left thalamus for language and verbal memory, and the right parietal lobe for visuoconstruction/-perception and visuospatial memory. This shows that our measure for PSCI represents multiple cognitive domains, and associated infarct locations are not driven by a single domain. Meanwhile, each cognitive domain was associated with a unique infarct pattern; for example, verbal memory was the only domain with a significant association with medial temporal lobe infarcts, which strengthens the face validity of this approach. The "viridis" color scale was used for visualization (based on https://cran.r-project.org/web/packages/viridis/index.html). FDR = False Discovery Rate, L = left, R = right.

Figure 4. Calibration plots for five-point location impact score as predictor for post-stroke cognitive impairment based on leave-one-cohort-out cross-validation

Calibration plots of predicted probabilities (x-axis) versus actual post-stroke cognitive impairment (PSCI) occurrence (y-axis) for the logistic regression model with the five-point location impact score as predictor. Each row indicates the cohort that was left out of the model derivation in the leave-one-cohort-out cross-validation and served as external validation sample. Cohorts are listed in descending order based on sample size. Distributions of actual 0 and 1 values are shown at the bottom of the graph; the *loess* smoother (with 95% confidence band) is shown in black; the ideal 45 degree line is shown in red. The actual outcomes, stratified according to risk groups (20% of validation sample per stratum), are indicated by triangles. Calibration and discrimination measures are shown in the upper left corner of the plots. The top row shows the original results, and the bottom row shows the results after adjusting for cohort-specific PSCI occurrence (i.e. to match the 44% in the total sample). For two cohorts (Hallym VCI and COAST) this adjustment was not needed because PSCI occurrence matched the overall occurrence (45% and 42% respectively). Two cohorts are not included in this figure (CODECS and Mild Stroke Study 2), because model validation was unsuccessful due to insufficient sample size or low PSCI occurrence.

Table 1. Demographics and clinical characteristics of individual cohorts and the total sample

	Bundang VCI (N=753)[23,24]	**CASPER (N=104)**[25]	**COAST (N=74)**[26]	**CODECS (N=27)**[12]	**CROMIS-2 (N=97)**[27]	**CU-STRIDE (N=410)**[11]	**GRECogVASC (N=316)**[10]	**Hallym VCI (N=641)**[23,24]	**MSS-2 (N=100)**[28]	**PROCRAS (N=177)**[29]	**STROKDEM (N=138)**[30]	**USCOG (N=113)**[21]	**All studies (n = 2950)**
Demographics													
Country of inclusion	Republic of Korea	NL	Singapore	NL	UK	Hong Kong	France	Republic of Korea	UK	NL	France	NL	
Ethnicity	Korean	Caucasian *	Singaporean Chinese (70%), Malay (22%), Indian (8%)	Caucasian	Caucasian**	Chinese	Caucasian	Korean	Caucasian	Caucasian	Caucasian	Caucasian	36% Caucasian; 47% Korean; 16% Chinese; 1% other
Age in years, mean (SD)	69·8 (10·8)	64·1 (10·8)	58·4 (10·5)	59·2 (16·5)	73·7 (9·2)	68·6 (10·4)	63·7 (10·6)	65·1 (11·9)	65·7 (11·5)	69·6 (9·4)	64·9 (12·1)	60·0 (14·9)	66·8 (11·6)
Female, n (%)	306 (40·6)	27 (26·0)	21 (28·4)	12 (44·4)	43 (44·3)	163 (39·8)	121 (38·3)	268 (41·8)	38 (38·0)	58 (32·8)	53 (38·4)	47 (41·6)	1157 (39·2)
Education category,[a] n (%)													
- Less than high school	402 (53·4)	42 (40·4)	52 (70·3)	8 (29·5)	6 (6·2)	381 (92·9)	233 (73·7)	351 (54·8)	19 (19·0)	79 (44·6)	86 (62·3)	44 (38·9)	1703 (57·7)
- High school	146 (19·4)	19 (18·3)	14 (18·9)	4 (14·8)	78 (80·4)	5 (1·2)	30 (9·5)	154 (24·0)	62 (62·0)	53 (29·9)	16 (11·6)	27 (23·9)	608 (20·6)
- Technical/college	37 (4·9)	35 (33·7)	7 (9·5)	5 (18·5)	N/A	8 (2·0)	35 (11·1)	31 (4·8)	N/A	39 (22·0)	11 (8·0)	26 (23·0)	234 (7·9)
- University or higher	168 (22·3)	8 (7·7)	1 (1·4)	10 (37·0)	13 (13·4)	16 (3·9)	18 (5·7)	105 (16·4)	19 (19·0)	6 (3·4)	25 (18·1)	16 (14·2)	405 (13·7)
Clinical characteristics													
NIHSS baseline, median (IQR)	3 (2-5)	N/A	3 (1-7)	0 (0)	3 (2-6)	4 (2-6)	3 (1-6)	2 (1-4)**	1 (0-2)	3 (2-4)	0 (0-1)	N/A	3 (1-5)***
IQCODE, median (IQR)	3·3 (3·1-3·7)**	3·1 (3·0-3·3)**	3·0 (3·0-3·1)	N/A	3·0 (3·0-3·3)	N/A	0% impaired	3·1 (3·0-3·3)***	N/A	3·0 (3·0-3·1)**	3·0 (3·0-3·1)	N/A	3·1 (3·0-3·4)***
Cognitive assessment timing, n days after event, median (IQR)	104 (10-170)	87 (81-99)	121 (105-152)	90 (N/A)	4 (2-9)	154 (129-176)	178 (161-186)	98 (90-105)	142 (53-383)	35 (29-40)	189 (178-199)	6 (4-9)	105 (74-170)*
Clinical history of stroke	104 (13·8)	5 (4·8)	10 (13·5)[b]	0 (0·0)	6 (6·2)*	50 (12·2)	22 (7·0)	85 (13·5)**	10 (10·0)	23 (13·0)	12 (8·7)	0 (0·0)	327 (11·1)*
Brain imaging													
Scan sequence/modality used for infarct segmentation	DWI (97%), FLAIR (3%)	FLAIR	DWI (41%), T2 (5%), CT (54%)	DWI (19%), FLAIR (63%), CT (19%)	DWI	DWI (75%), CT (25%)	T1	DWI (98%), FLAIR (2%)	FLAIR	FLAIR	DWI	FLAIR (34%), CT (66%)	DWI (66%), T2/FLAIR (16%), CT (8%), T1 (11%)
Normalized acute infarct volume in ml, median (IQR)	3·8 (1·2-16·5)	3·4 (0·9-13·2)	6·8 (2·0-32·6)	10·3 (1·1-26·2)	4·5 (1·5-16·1)	2·3 (0·9-12·9)	1·3 (0·3-5·7)	2·0 (0·9-11·3)	2·6 (1·2-11·3)	4·4 (1·3-21·0)	1·6 (0·6-8·6)	19·6 (3·5-51·9)	2·7 (1·0-14·1)
Imaging timing, n days after event, median (IQR)	5 (4-6)	87 (81-99)	2 (1-4)	34 (5-98)**	5 (3-9)	1 (0-2)	178 (161-186)	1 (1-2)	4 (2-9)	33 (27-40)	3 (3-3)	5 (3-8)	4 (1-9)*

* Missing in <1%; ** Missing in 1-10%; *** Missing in >10%; [a] Education categorization per cohort is shown in the appendix (appendix, p.28). [b] Combined variable for stroke and/or TIA.

Abbreviations: CT, computed tomography; DWI, diffusion-weighted imaging; FLAIR, fluid attenuated inversion recovery; IQR, interquartile range; NL, the Netherlands; SD, standard deviation; TIA, transient ischemic attack; UK, United Kingdom

Table 2. Univariate logistic regression results for the continuous and five-point location impact score in the total sample (N=2950)

	OR per point (95% confidence interval)	**Predicted probability PSCI[a]**
Location impact score – continuous (max. range: -1·25 to 2·44)	2·36 (2·06-2·70)***	P5 (score=-0·21): 25% P50 (score=0·68): 43% P95 (score=1·77): 65%
Location impact score – five-point (reference = 1, 0-20th percentile; includes cont. score range -1·25 to 0·25)		30%
2 (20-40th percentile; includes cont. score range 0·25 to 0·54)	1·30 (1·02-1·66)*	36%
3 (40-60th percentile; includes cont. score range 0·54 to 0·83)	1·48 (1·16-1·88)**	39%
4 (60-80th percentile; includes cont. score range 0·83 to 1·22)	2·08 (1·64-2·65)***	47%
5 (80-100th percentile; includes cont. score range 1·22 to 2·44)	4·40 (3·45-5·62)***	66%

Statistically significant variables: *p<0·05; **p<0·01; ***p<0·001. [a] Compared to overall PSCI occurrence of 44%. For the continuous score, probabilities for the 5th/50th/95th percentiles are shown.. OR = odds ratio, PSCI = post-stroke cognitive impairment. The visual rating scale of the location impact score is available online: https://metavcimap.org/features/software-tools/location-impact-score/