Right hemispheric white matter hyperintensities improve the prediction of spatial neglect severity in acute stroke

Lisa Röhrig [a], Christoph Sperber [a], Leonardo Bonilha [b], Christopher Rorden [c], Hans-Otto Karnath [a,c,*]

[a] *Division of Neuropsychology, Center of Neurology, Hertie-Institute for Clinical Brain Research, University of Tübingen, Tübingen 72076, Germany*
[b] *Department of Neurology, Emory University, Atlanta, GA 30322, USA*
[c] *Department of Psychology, University of South Carolina, Columbia, SC 29208, USA*

ARTICLE INFO

ABSTRACT

White matter hyperintensities (WMH) are frequently observed in brain scans of elderly people. They are associated with an increased risk of stroke, cognitive decline, and dementia. However, it is unknown yet if measures of WMH provide information that improve the understanding of poststroke outcome compared to only state-of-the-art stereotaxic structural lesion data.

We implemented high-dimensional machine learning models, based on support vector regression, to predict the severity of spatial neglect in 103 acute right hemispheric stroke patients.

We found that (1) the additional information of right hemispheric or bilateral voxel-based topographic WMH extent indeed yielded a significant improvement in predicting acute neglect severity (compared to the voxel-based stroke lesion map alone). (2) Periventricular WMH appeared more relevant for prediction than deep subcortical WMH. (3) Among different measures of WMH, voxel-based maps as measures of topographic extent allowed more accurate predictions compared to the use of traditional ordinally assessed visual rating scales (Fazekas-scale, Cardiovascular Health Study-scale).

In summary, topographic WMH appear to be a valuable clinical imaging biomarker for predicting the severity of cognitive deficits and bears great potential for rehabilitation guidance of acute stroke patients.

1. Introduction

White matter hyperintensities (WMH), also called white matter lesions or leukoaraiosis, are frequently seen as hyperintense areas on T2-weighted brain MRI of predominantly elderly people (Manolio et al., 1994; for a review, see Tang et al., 2021). While not fully understood, both vascular and non-vascular causes are involved (Schmidt et al., 2011). WMH of presumed vascular origin are seen to represent features of the cerebral small vessel disease (Wardlaw et al., 2013a; Wardlaw et al., 2013b). Because WMH increase with age (Manolio et al., 1994; Wardlaw et al., 2015; Wei et al., 2019), they were initially thought to be a consequence of the normal ageing brain. Recently, WMH have received more attention, as they are associated with an increased risk of stroke (Debette and Markus, 2010), stroke recurrence (Andersen et al., 2017; Georgakis et al., 2019), mortality (Debette and Markus, 2010; Georgakis et al., 2019), disturbed mobility (Guttmann et al., 2000), cognitive decline (Qi et al., 2019; Valdés Hernández et al., 2013), and dementia (Debette and Markus, 2010; Georgakis et al., 2019). Moreover, WMH are reported to be specifically related to recovery of cognitive but not other poststroke impairments (Hawe et al., 2018; Khan et al., 2019). WMH are also linked to cognitive deficits, as studies find associations of WMH extent with worse cognitive functions in Parkinson's disease (Huang et al., 2020; Linortner et al., 2020; Liu et al., 2021), as well as severity and rate of progression of primary progressive aphasia (Meier et al., 2020; Odolil et al., 2020). Other research reveals a connection between WMH severity and performance of aphasia in the chronic phase of stroke (Basilakos et al., 2019; Wilmskoetter et al., 2019). In sum, WMH seem to be a strong indicator of compromised brain health and brain reserve.

Table 1
Demographic and clinical data of the patient sample.

Parameter	Total (N = 103)	Neglect (N = 27)	No neglect (N = 76)	p
Age *(years)*	58.2 (13.7) [26–82]	60.5 (15.1) [27–82]	57.4 (13.2) [26–80]	0.354[a]
Sex *(F, M)*	37, 66	10, 17	27, 49	0.888[b]
Stroke to imaging *(days)*	3.5 (3.9) [0–16]	4.7 (4.7) [0–15]	3.1 (3.5) [0–16]	0.134[a]
Etiology *(I, H)*	96, 7	26, 1	70, 6	0.457[b]
Lesion volume *(cm³)*	27.5 (32.9) [0.1–182.0]	55.6 (44.4) [0.4–182.0]	17.5 (20.1) [0.1–98.9]	**< 0.001**[a]
Stroke to assessment *(days)*	3.4 (2.9) [0–14]	3.0 (2.3) [0–8]	3.5 (3.0) [0–14]	0.393[a]
Letter CoC	0.10 (0.23) [-0.06–0.96]	0.39 (0.30) [0.01–0.96]	0.01 (0.02) [-0.06–0.07]	**< 0.001**[a]
Bells CoC	0.13 (0.24) [-0.03–0.92]	0.46 (0.27) [0.12–0.92]	0.02 (0.03) [-0.03–0.14]	**< 0.001**[a]
Mean CoC	0.12 (0.23) [-0.01–0.91]	0.43 (0.28) [0.08–0.91]	0.01 (0.02) [-0.01–0.06]	**< 0.001**[a]
Visual field defects *(N)*	12	3	9	0.919[b]

Patients with mean CoC $\geq$ 0.082 were defined as having spatial neglect (cut offs for the letter and bells cancellation tasks are 0.083 and 0.081, respectively (Rorden and Karnath, 2010)). Results are given in either mean (standard deviation) [range] or number of patients. 'Lesion volume' represents the normalized volume in MNI space. *p*-values < 0.05 were considered significant and are highlighted in bold type.
Abbreviations: *F* – females, *M* – males, *I* – Ischemia, *H* – Hemorrhage.
[a] Unequal variances *t*-test (continuous measure).
[b] Chi²-test (categorical measure).

Only a few studies have investigated whether the extent of WMH is also a predictor of the occurrence and severity of cognitive impairment in the acute phase of stroke. Bahrainwala and colleagues (2014) reported an association between WMH severity (adjusted for age, sex, race, and infarct volume) and the probability of having spatial neglect after an acute right hemispheric stroke. Stroke patients with spatial neglect show spontaneous and sustained deviation of their eyes and head as well as focus of attention toward the ipsilesional, typically right side of space (Karnath and Dieterich, 2006; Karnath and Rorden, 2012). Simultaneously, they fail to observe and react to stimuli, objects, and people on the contralesional side. Bahrainwala and colleagues (2014) found that for the same size of stroke lesion, a greater extent of WMH increased the risk of suffering from any and also more severe spatial neglect in the acute phase of stroke. Of high interest, the authors provided evidence that prestroke WMH affect poststroke acute attentional impairments. Such biomarkers could be used in the guidance and planning of poststroke rehabilitation and care.

Unfortunately, this seminal study used data with reduced information content as predictors, namely infarct volume without localization-related information and the extent of WMH as an ordinal scaled measure. However, voxel-wise representation is commonly used in state-of-the-art prediction algorithms (Salvalaggio et al., 2020; Siegel et al., 2016), and previous research demonstrated that voxel-based maps are better able to describe the brain's complexity compared to lesion volume (Rondina et al., 2016) and are thus likely to achieve more accurate behavioral predictions. The ordinal clinical visual rating scales used for the WMH assessment bear also the risk of producing subjective discrepancies between raters. Moreover, Bahrainwala and colleagues (2014) predicted dichotomized variables, i.e., the presence versus absence of spatial neglect, despite the fact that symptom severity varies between patients (Rorden and Karnath, 2010). Furthermore, the authors did not test whether the localization of WMH has any significance for the severity of acute spatial neglect. Specifically, are left-sided, right-sided, and bilateral WMH equally predictive, and is it relevant whether the WMH are located near the lateral ventricles or within the deep subcortical white matter?

The first aim of the present study was to investigate whether the precise, voxel-based measurement of WMH extent by using semi-automated MRI analysis improves the accuracy of predicting neglect severity in acute stroke patients, compared to using rating-based

measures. We tested the visual rating scales by the Cardiovascular Health Study (CHS; Manolio et al., 1994), which was applied by Bahrainwala and colleagues (2014), and Fazekas (Fazekas et al., 1987), which is the most commonly used WMH rating scale (Georgakis et al., 2019). Furthermore, instead of determining only the presence versus absence of spatial neglect, we analyzed the deficit as a continuous variable. We thus applied machine learning-based support vector regression (SVR) as it allows to predict continuous variables with high-dimensional features. Second, we were interested in examining the possible impact of WMH localization within the hemisphere(s) on predicting the severity of neglect. For all models, we used the stroke lesion maps as a baseline. While a plethora of factors – including microstructural damage, age, sex, education, and brain atrophy – impact on the occurrence of poststroke deficits, they only do so in an interaction with the brain lesion.

2. Materials and methods

2.1. Participants

We included 103 individuals with right hemisphere stroke who were consecutively admitted to the Center of Neurology of Tübingen University; for demographic and clinical data, see Table 1. The patients had participated in previous studies of our group. Inclusion criteria for the current investigation were first-ever, right hemispheric stroke with a demarcated lesion and available MRI scans including a FLAIR (fluid-attenuated inversion recovery) image. Patients were not included when the period between stroke onset and imaging or neuropsychological examination was longer than 16 days. Patients with bilateral strokes or brain tumors were also not included. All patients gave their informed consent for participation and scientific data usage. The study was conducted in accordance with the revised Declaration of Helsinki and was approved by the ethics committee of the medical faculty of Tübingen University.

2.2. Behavioral data

Patients were tested for spatial neglect using the letter cancellation test (Weintraub et al., 1985) and bells cancellation test (Gauthier et al., 1989) as standard paper-and-pencil tests. We calculated the *Center of Cancellation* (CoC; Rorden and Karnath, 2010) as a continuous score

ranging between -1 (right-sided neglect) and +1 (left-sided neglect). The CoC is a sensitive measure capturing both the number of omissions as well as their location. For each patient, the mean CoC of both cancellation tasks was used for the analyses. If one of both test results was not available ($N = 5$), we used the remaining CoC score. Primary visual field defects (hemi- or quadrantanopia) were assessed via usual neurological confrontation testing.

2.3. Imaging data and analysis

We used the axial MR images acquired during the patients' clinical investigation at the Center of Neurology Tübingen. Since WMH are most visible on FLAIR scans (Wardlaw et al., 2015), we used these for the assessment of WMH extent. For quantification of the stroke lesion, we used co-registered diffusion-weighted imaging (DWI) scans if imaging was acquired within 48 h after stroke onset ($N = 47$); otherwise, we used the FLAIR scans again ($N = 56$).

For lesion volume quantification, we applied the following steps performed by the use of SPM 12 (*Statistical Parametric Mapping*; Wellcome Department of Imaging Neuroscience, London, UK) for MATLAB R2019a (The MathWorks, Inc., Natick, USA). Stroke lesion delineation was conducted semi-automatically, using the 'Clusterize Toolbox' (Clas et al., 2012; de Haan et al., 2015) implemented in SPM (see Fig. 1A). Clusterize is an algorithm automatically marking hyperintense (or hypointense) areas according to an intensity threshold; we used the default minimum cluster size of 100 voxels (Clas et al., 2012). A voxel-based binary brain map was generated by selecting the clusters belonging to the lesion for each image slice separately by one investigator (LR) who was blind to the behavioral measures. The resulting lesion map was then normalized (together with the respective anatomical image) to MNI (*Montreal Neurological Institute*) space (Collins et al., 1994) with voxels of size 1 mm^3 by using the 'Clinical Toolbox' (Rorden et al., 2012) for SPM and its age-matched templates (this step leveraged the high-resolution T1-weighted scan, if available). Lesions were controlled by either cost-function masking or enantiomorphic correction (cf. Karnath et al., 2019) depending on the available scan modalities and normalization outcome. Lateral ventricles, cerebellum and extracerebral space were masked to address normalization inaccuracies.

The extent of WMH was first determined quantitatively by directly obtaining each patient's volumetric WMH map. For delineation of the WMH map, we used the same semi-automatic procedure via the 'Clusterize Toolbox' as for the patients' volumetric stroke lesion map (see above and Fig. 1A), except that we now set the minimum cluster size to 10 voxels because WMH may consist of small punctate foci. WMH volume maps were normalized together with the same patient-specific spatial transformation parameters as for the patient's stroke lesion normalization step. To exclude any potential overlap between both stroke and WMH maps in cases where WMH were adjacent to the stroke lesion, the WMH map was masked by the respective lesion map for each patient. For investigating laterality effects, we defined left and right WMH maps by separating the overall normalized WMH map at the midsagittal plane.

Second, the extent of WMH was also assessed visually by using two visual rating scales: (1) the scale described in the *Cardiovascular Health Study* (CHS; Manolio et al., 1994) applied by Bahrainwala and colleagues (2014) and (2) the more frequently used rating scale by Fazekas and colleagues (1987). Both scales are based on the examiner's subjective visual inspection of the original FLAIR images. The CHS visual rating scale is based on 8 images depicting increasing severity stages of WMH; the scale ranges from 0 to 9, where 9 is defined as most severe. The Fazekas rating scale assesses periventricular white matter hyperintensities (PV-WMH) and deep subcortical white matter hyperintensities (DS-WMH) separately. Ratings are given as follows: PV-WMH: 0 – absent, 1 – 'caps' or pencil-thin lining, 2 – smooth 'halo', 3 – irregular PV-WMH extending into the deep white matter; DS-WMH: 0 – absent or a single punctate focus, 1 – multiple punctate foci, 2 –

beginning confluency of foci, 3 – large confluent areas. The total Fazekas score is calculated by summing up PV-WMH and DS-WMH scores and therefore ranges from 0 to 6, where 6 is defined as most severe. One author (LR) conducted the ratings of all patients that were used in the analyses without knowing any clinical or neuropsychological data and diagnoses. To check for interrater reliability, another author (HOK) rated a subset of 20 examples; again without knowing any clinical or neuropsychological data and diagnoses. For each rating scale, we calculated Cohen's linearly weighted kappa (κ) with the relative agreement to estimate consensus and Spearman correlation (r_s) to estimate consistency (Stemler, 2004): CHS-scale – $\kappa = 0.72$ (90.7 %) and $r_s = 0.87$; Fazekas total score – $\kappa = 0.40$ (82.5 %) and $r_s = 0.85$; Fazekas PV-WMH score – $\kappa = 0.58$ (88.3 %) and $r_s = 0.76$; Fazekas DS-WMH score – $\kappa = 0.25$ (73.3 %) and $r_s = 0.71$. Although the weighted kappa for the DS-WMH score represents only fair consensus in rating (Landis and Koch, 1977), the corresponding strong correlation demonstrates high consistency, which is sufficient for a reliable regression analysis.

2.4. Statistical analysis

2.4.1. Prediction of spatial neglect severity

To investigate whether WMH extent information – in addition to the acute stroke lesion – increases prediction accuracy of the severity of spatial neglect, we applied the supervised learning algorithm support vector regression (SVR; Brereton and Lloyd, 2010; Smola and Schölkopf, 2004) that can predict a continuous variable. SVR was implemented using the 'libsvm' package (Chang and Lin, 2011) for MATLAB via custom scripts. The feature matrix consisted of the lesion map and – in case of combined models (see below for details) – the WMH data of all patients. Matrix rows comprised patients, while columns contained features of the predictive model: vectorized volume map(s) including the dichotomized status (damaged or not) of all voxels, and optionally also the WMH rating score. The vector of the mean CoC scores, reflecting the severity of spatial neglect, served as the dependent variable. Since for our sample of only right hemispheric stroke patients no right-sided neglect was expected, negative mean CoC values ($N = 13$ with a *mean original CoC* = -0.006) were set to 0 to avoid difficulties in interpreting the directionality of predicting a bi-polar target variable. Moreover, we square root-transformed the mean CoC values to obtain less skewed data (see Fig. S1 in the Supplement) as implemented in previous research (Kasties et al., 2021).

In a pilot investigation, we found worse model performances when the ordinal WMH ratings were transformed into dummy variables. We therefore treated the WMH scores as continuous measures. We applied the mean normalization for the WMH scale values to obtain features between zero and one (i.e., the highest WMH rating score was assigned to the value of one, whereas the smallest WMH rating score was assigned to the value of zero). To focus only on the most relevant voxels of the lesion maps, we excluded voxels that were damaged in less than five patients in all models investigated. In contrast, we altered this criterion for the inclusion of voxels in the WMH maps. Ordinal ratings took the wholistic WMH pattern into account, including brain regions that were only rarely affected by WMH, such as small foci of DS-WMH. Since we aimed to directly compare rating-based with voxel-based WMH measures, we included all WMH map voxels that were damaged at least once to make rating-based and topographic data comparable. In addition, we applied a principal component analysis (PCA) to reduce the high-dimensional feature space. This feature reduction procedure was shown to work best for high-dimensional data by yielding most accurate model fits (Kasties et al., 2021). The PCA was only applied to voxel-based matrices (lesion map, WMH map) without the inclusion of single-value variables (WMH volume, scale ratings). We kept and normalized principal components that cumulatively explained at least 94 % of the total variance: for lesion maps 47 (out of 102), for bilateral WMH maps 77 (i.e., 124 when concatenated with components of lesion maps), for added matrices of lesion and bilateral/left/right WMH maps
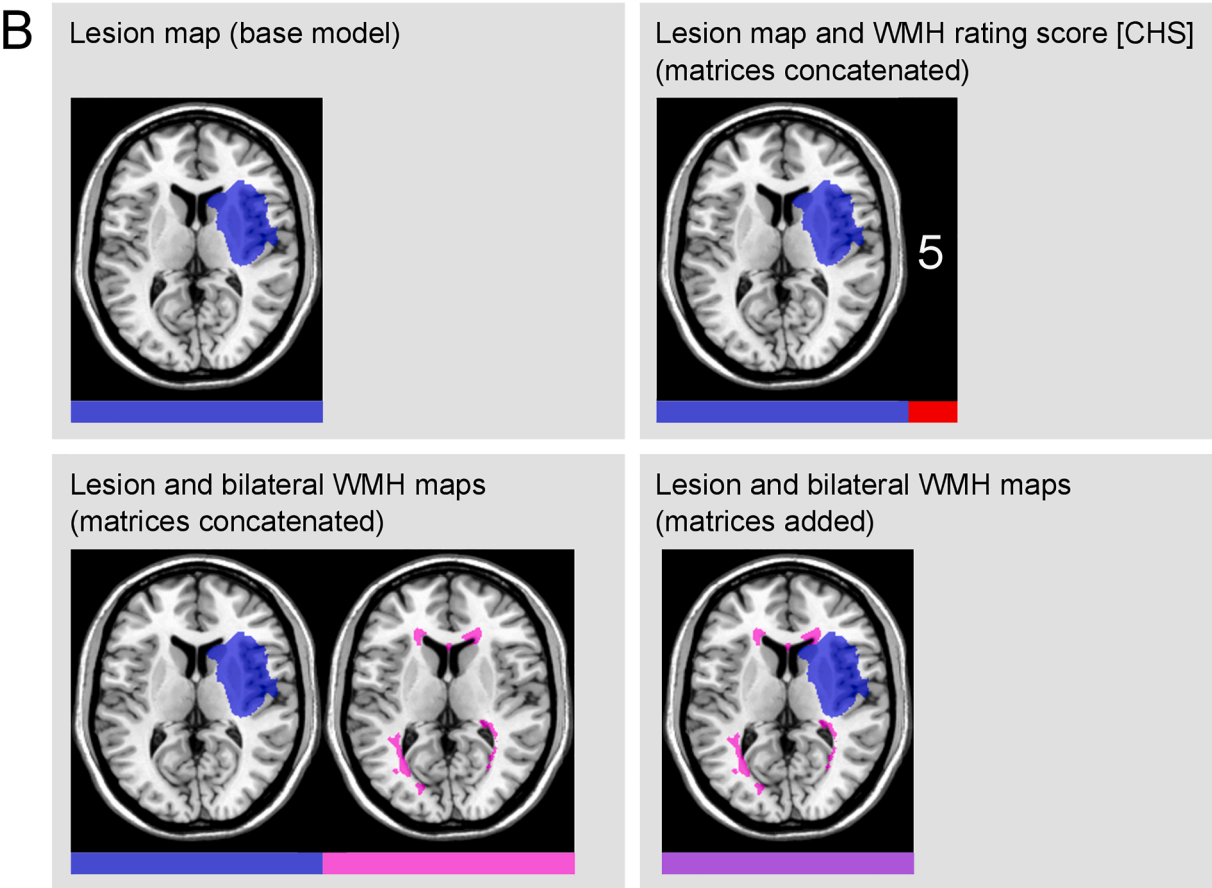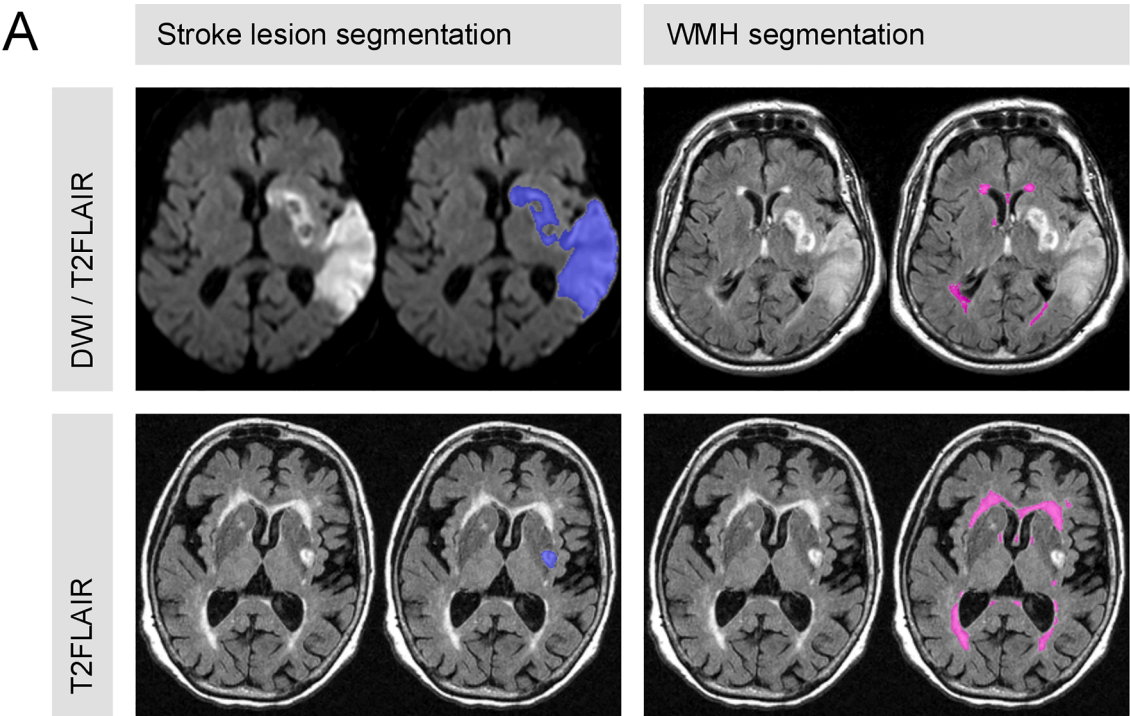
(caption on next page)

Fig. 1. Lesion segmentation and model versions. (A) Segmentations of stroke lesion (blue) and white matter hyperintensities (WMH; pink) are depicted for two exemplary patients. Either DWI and T2FLAIR or only T2FLAIR images were used depending on stroke onset. Maps were generated using the semi-automated 'Clusterize Toolbox' (de Haan et al., 2015). **(B)** Depending on the model, normalized lesion map (blue) and WMH map (pink) of an exemplary patient are presented on the ch2-template via MRIcron (Rorden and Brett, 2000). These matrices of damaged voxels were either attached to each other (concatenated) or matrices were summed up (added). In case of concatenation of lesion and WMH maps, we treated both variants of voxel-information as separate features. The matrix addition approach can be roughly explained as the simultaneous delineation of lesion and WMH in an overall brain damage map, as the expansion of the lesion map by the extent of WMH, or as the treatment of WMH as part of the lesion itself. For the WMH rating (CHS-scale exemplary), a single value representing the WMH severity was attached to the vectorized matrix of lesioned voxels (concatenated; for WMH volume equivalent). When concatenating both periventricular and deep subcortical WMH ratings (Fazekas-scale), two values were appended. Color bars illustrate the corresponding feature matrix (for one patient) that differed among the model versions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Nested cross validation
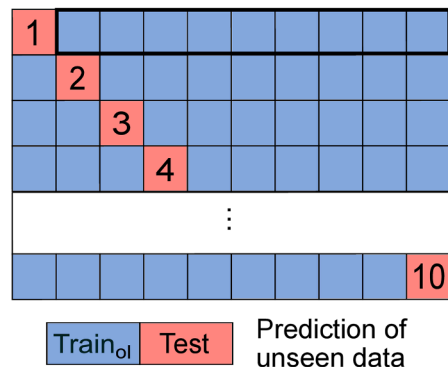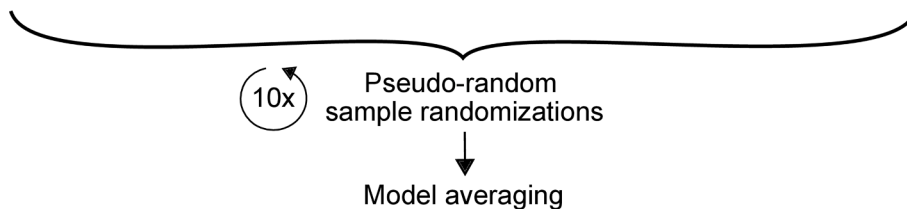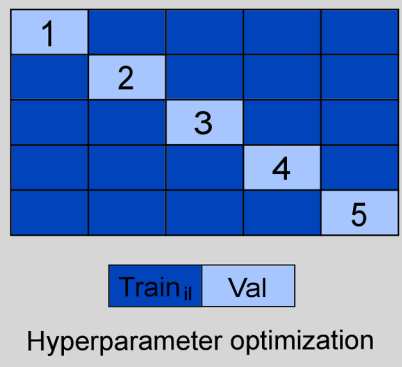
Fig. 2. Overview of the applied nested cross-validation procedure. Nested cross-validation with 10-fold outer loop and fivefold inner loop is illustrated. In the outer loop, nine folds served as training set ('$Train_{ol}$'), which is passed to the inner loop. In the inner loop, four folds served as training set ('$Train_{il}$'), with which the hyperparameters C and γ were trained and tested on the remaining fifth fold (validation set, 'Val'). This optimization process aimed to find the model that minimizes the mean absolute error (MAE). The model with the winning C-γ-combination of the inner loop was then trained on the whole training set of the outer loop and tested on the hold-out 10th test fold. The model fit was calculated based on the predictions made during the outer loop. We used a pseudo-random number stream to shuffle the sample order and repeated the whole algorithm 10 times with different sample randomizations to include variance-driven effects. In the end, predictions of the 10 model repetitions were averaged to get the averaged model and to calculate the final model fit.

73/65/66 components, respectively, indicating that WMH maps were more high-dimensional than lesion maps. PCA-results are visualized in Fig. S2. The threshold of 94 % was used because a test of manually selected cut-offs for the lesion maps revealed it as the optimal number of components. Results obtained when setting the threshold to 98 % are reported in our preprint (https://doi.org/10.1101/2022.04.08.22273547).

We implemented an epsilon-SVR with a radial basis function (RBF) kernel, where the default value for epsilon ($\varepsilon = 0.1$) was used in accordance with previous lesion-behavior modeling studies (Wiesen et al., 2019; Zhang et al., 2014). Previous studies investigating lesion-behavior relationships have reported better performance of nonlinear kernels compared to linear ones (Hope et al., 2018; Zhang et al., 2014). To avoid overfitting and to promote generalization of the fitted model, we implemented a nested cross-validation (CV) procedure (Krstajic et al., 2014; Varoquaux et al., 2017) (Fig. 2). In the outer loop, we performed a 10-fold CV. Folds were almost equally sized. In one iteration, nine folds of the patient sample were assigned to the training set and the remaining 10th fold served as a test set. The training set of the outer loop was passed to the inner loop and contained all training- and validation folds of the inner CV. In the inner loop, we applied a fivefold CV, i.e., four folds served as training set of the inner CV and the remaining fifth fold represented the validation set. During the inner loop, the hyperparameters C and γ were optimized. For that, a grid search was implemented ($C = 2^{(-5)}, 2^{(-4)}, ..., 2^{15}$, and $\gamma = 2^{(-15)}, 2^{(-14)}, ..., 2^5$). Each combination of C and γ was trained on the training sample and tested on the remaining validation fold. We defined the *Mean*

Absolute Error (MAE) as our parameter of interest. After the fivefold CV (inner loop), the aforesaid parameter was averaged for that specific C-γ-combination. This procedure was repeated for each parameter combination. The algorithm was designed to minimize the MAE: the C-γ-combination that yielded the smallest mean MAE was considered to be the best model generated during the inner loop. The model with the winning hyperparameter-combination was then re-trained on the whole training set of the outer loop and tested on the hold-out test set. At the end of the whole nested CV, the outcome of each patient was predicted once during the outer loop.

Since the resulting model fit is dependent on the patient assignments to training-, validation- and test sets, we aimed to overcome that issue of variance by implementing model averaging (Arlot and Celisse, 2010; Varoquaux et al., 2017). To get a more generalized model fit, we repeated the whole model fitting procedure 10 times (cf. Fig. 2). For each model repetition, we applied a different sample randomization. For reproducibility and comparability issues, we used a specified random number stream. The sample was pseudo-randomized for each model repetition according to a sub-stream of the applied random number generator, so that the randomization of patients of a particular model repetition was the same across model versions. The predictions of all 10 model repetitions were then averaged for each patient. These final predictions were used to calculate the overall *Coefficient of Determination* (R^2; for formula, see, e.g., Chicco et al., 2021), which represents the proportion of variance explained. The R^2 served as the model fit parameter when comparing the investigated SVR model versions as previously recommended (Chicco et al., 2021).

The feature matrix of the base model consisted of the voxel-based vectorized binary lesion map. For the other SVR models, we had to connect different types of data into one feature matrix. We used two methods of combining lesion and WMH: concatenation and addition of matrices (Fig. 1B). (1) We concatenated the matrix of lesion maps and the vector of WMH scale values to get one enlarged matrix, separately for the CHS-scale, Fazekas PV-WMH score, Fazekas DS-WMH score, and for both PV-WMH and DS-WMH scores concatenated. (2) We also merged the lesion matrix with either the vector of WMH volumes or the matrix of WMH maps into one concatenated large matrix to directly compare the effect of using volumetric information instead of a single ordinal rating. (3) Further, we performed a voxel-wise matrix addition to join the lesion and WMH maps together to one voxel-based volume. (For that, we kept for both volumetric maps all voxels that are lesioned at least five times among the lesion maps and/or at least once among the WMH maps. The reduced lesion and WMH maps were added afterwards.) For the investigation of laterality effects, we also calculated models with either a left or a right hemispheric WMH map; here, we used the matrix fusion technique (matrix concatenation or addition) that revealed the best model performance for using the bilateral WMH extent.

2.4.2. Permutation tests

To validate whether a numerical increase in model fit is significantly above chance, we applied nonparametric Monte-Carlo simulation (permutation) testing to all models that yielded a better performance compared to the base model (the model that used only lesion data). For that, the model was calculated with a shuffled patient assignment of WMH data to investigate whether the patient-specific lesion-WMH-combination is relevant for obtaining an improved prediction performance. This way, we checked if WMH extent improves prediction significantly. In case of matrix addition, we controlled for any overlaps of lesion and WMH maps that may have occurred by setting all non-zeros to ones (keeping matrix binary). For each of the 10 model repetitions with different patient randomizations, we applied 5,000 permutation iterations. The predictions of all 5,000 permutations were averaged across the 10 model repetitions. For each of the 5,000 averaged models with permuted WMH severity, the *Coefficient of Determination* (R^2) was then calculated. From the distribution of 5,000 R^2 obtained by the permuted variable (i.e., WMH data), one can conclude if the true WMH data resulted in a model performance significantly better than chance: if the 95 %-best R^2 is smaller than the true R^2, then the true R^2 is significant on an alpha-level of 0.05. Using the predicted and true outcome scores, the *p*-value was calculated additionally. Since we were only interested in an improvement of prediction accuracy, we used a one-tailed significance test.

Again, for reproducibility issues, the random number stream that was implemented for the cross-validation procedure was also applied for the permutation process (5,000 sub-streams were used to permute the WMH data pseudo-randomly). The use of pseudo-random permutations had the further advantage that we were able to directly compare the permutation tests of different model versions and their repetitions, and to convincingly average the predictions that resulted from the 10 permutation repetitions.

3. Results

3.1. Clinical and anatomical aspects

Fig. 3 illustrates the topography of simple overlaps of acute lesions and WMH in the group of right hemispheric stroke patients with and without spatial neglect ($N = 27$ and $N = 76$, respectively); Table 2 gives the results of the different measures of WMH for the two groups. A first result was that our continuously admitted sample of 103 acute stroke patients did not include patients without any WMH (i.e., CHS score of 0, WMH volume of 0 ml). With respect to the extent of WMH (regardless of
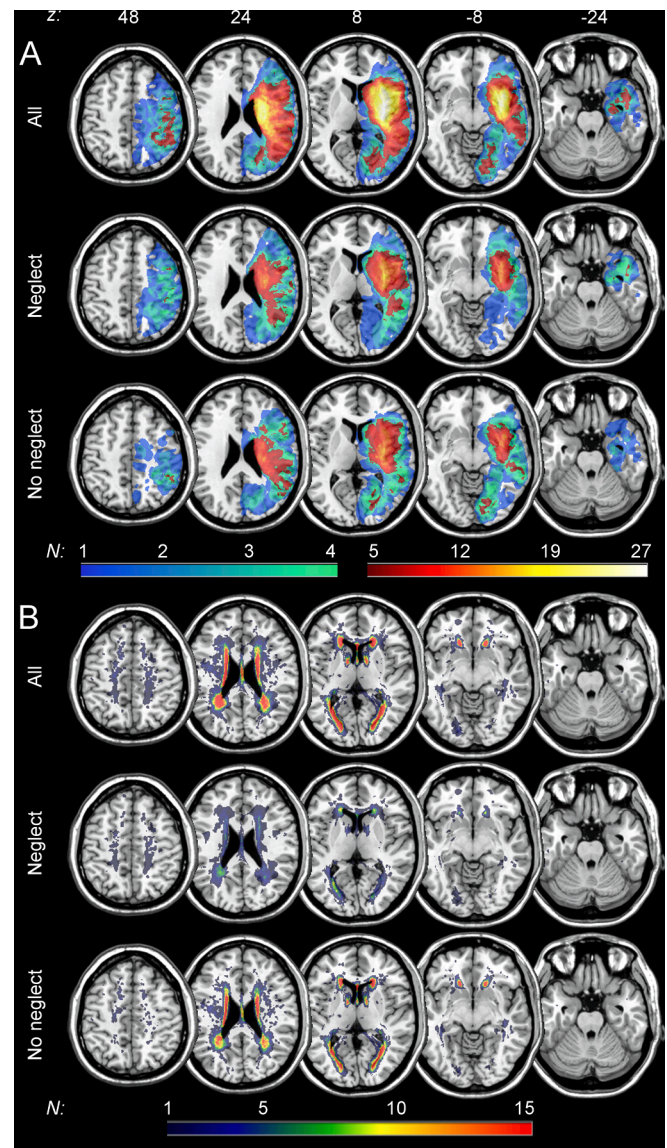
Fig. 3. Simple overlay plots. Overlaps of (**A**) normalized acute lesions and (**B**) normalized white matter hyperintensities (WMH) are shown for all patients ($N = 103$) and for patients with or without spatial neglect ($N = 27$ or $N = 76$, respectively) on the ch2-template in MNI space via MRIcron (Rorden and Brett, 2000). Displayed axial slices refer to z-coordinates 48, 24, 8, -8 and -24 mm. The color of the voxels represents the number of patients with damage to this voxel ($N_{min} = 1$; $N_{max} = 27$ or $N_{max} = 15$, respectively). For lesions (**A**), cold colors refer to voxels that were damaged in less than five patients, whereas voxels displayed in warm colors were lesioned in at least five patients, i.e., warm-colored voxels shown for all patients were used for the prediction analyses. For WMH (**B**), all voxels damaged in at least one patient were used for prediction.

whether measured ordinally or volumetrically), patients suffering from spatial neglect did not differ significantly from patients without the disorder (all $p > 0.05$; Table 2). However, we observed numerically larger WMH volumes in patients with spatial neglect (cf. Table 2). Spearman correlations (Table S1 in the Supplement) revealed highly significant positive relationships between all WMH measures ($0.36 \leq r_s \leq 0.89$; all $p < 0.001$). The three variables CHS score, total Fazekas score and overall WMH volume (single value) were very strongly correlated (all $r_s \geq 0.84$). Contrary, PV-WMH and DS-WMH showed only a weak correlation ($r_s = 0.36$). Moreover, we observed the previously described moderate positive relationship between age and any WMH measure

Table 2
WMH measures for neglect versus non-neglect patients.

Parameter	Total (N = 103)	Neglect (N = 27)	No neglect (N = 76)	p
CHS (0–9)	3 [1–9]	3 [1–9]	3 [1–8]	0.944[a]
Fazekas (0–6)	2 [1–6]	2 [1–6]	2 [1–6]	0.412[a]
PV-WMH (0–3)	1 [0–3]	1 [1–3]	1 [0–3]	0.901[a]
DS-WMH (0–3)	1 [0–3]	1 [0–3]	1 [0–3]	0.166[a]
Volume (cm³)	8.0 (8.9) [0.4–64.8]	9.6 (13.0) [1.2–64.8]	7.5 (6.9) [0.4–39.4]	0.432[b]

Results are given as median [range] for ordinal data (for total scale range see parentheses) or mean (standard deviation) [range]. 'Fazekas' represents the total score of the Fazekas scale (scores of PV-WMH and DS-WMH summed up). 'Volume' represents the normalized volume in MNI space. p-values < 0.05 were considered significant.

Abbreviations: *CHS* – Cardiovascular Health Study, *PV-WMH* – periventricular WMH, *DS-WMH* – deep subcortical WMH, *WMH* – white matter hyperintensities.

[a] Mann-Whitney *U* test (ordinal measure).
[b] Unequal variances *t*-test (continuous measure).

($0.39 \leq r_s \leq 0.53$; all $p < 0.001$); no such correlations were found between these WMH measures and neglect severity ($p > 0.05$).

3.2. Prediction of spatial neglect severity

Fig. 4 gives an overview of model performances, i.e., prediction accuracies of the averaged models as well as the performances of single model repetitions for all model versions; Fig. S3 and Fig. S4 show corresponding scatter plots. The averaged base model that used only lesion maps (no additional WMH data) in the feature matrix revealed a prediction accuracy of $R^2 = 0.368$. Looking at performances of the single model repetitions (Fig. 4), models using the CHS-scale value, both Fazekas scores (PV-WMH and DS-WMH) concatenated, or WMH volume showed variance in prediction accuracy depending on the sample randomization ($R^2_{max} - R^2_{min} \geq 0.101$). Compared to that, the other models seemed to be more stable in model fit across repetitions ($R^2_{max} - R^2_{min} \leq 0.097$).

3.2.1. Rating-based WMH measures

When attaching either the CHS-scale values, DS-WMH scores, or concatenated PV-WMH and DS-WMH scores to the lesion matrix, the averaged model yielded a poorer performance compared to the base model. The model using lesion map and PV-WMH score achieved a slightly improved accuracy and thus the best model performance among all WMH ratings, indicating that PV-WMH was more predictive than DS-
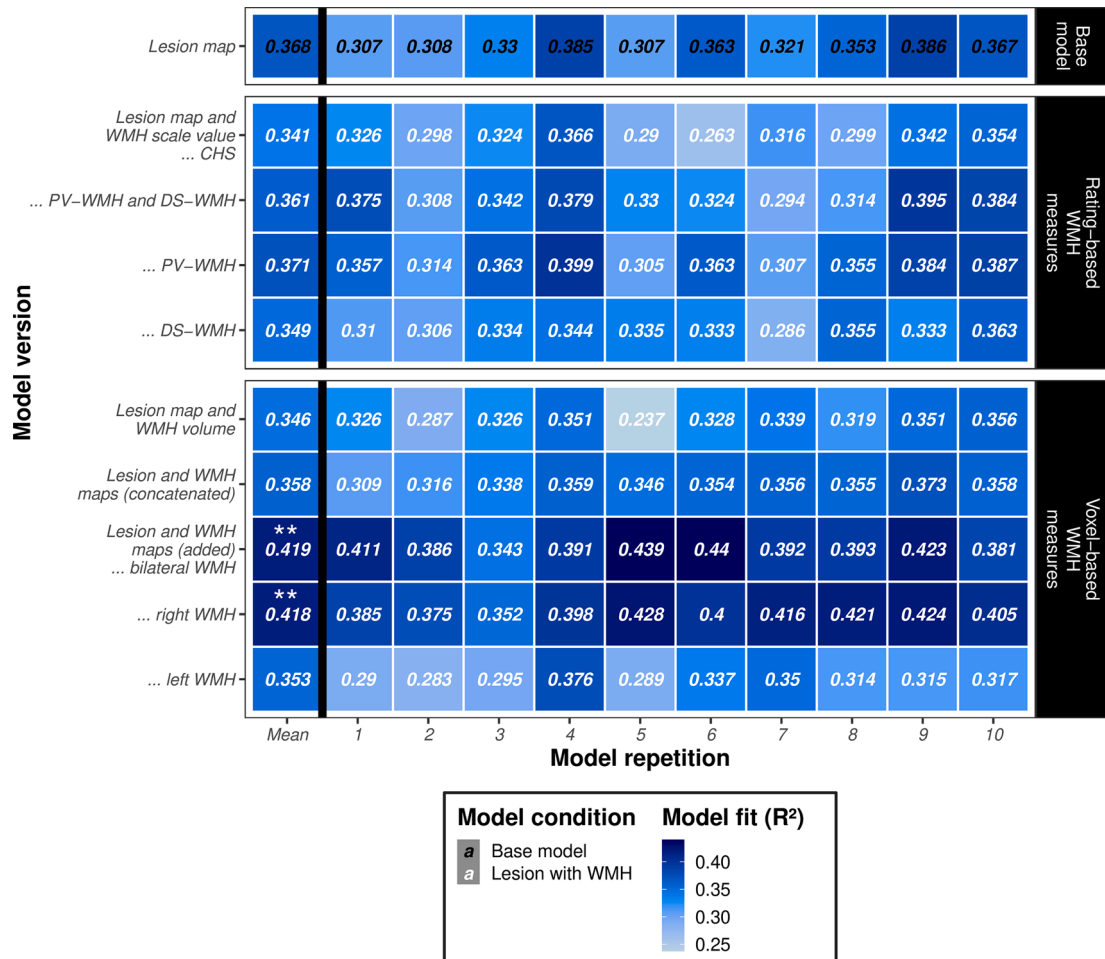
Fig. 4. Model performances. Heatmap illustrating accuracies of predicting neglect severity (CoC score; see Rorden and Karnath, 2010) for the cross-validated support vector regression (SVR) models. Prediction accuracy was measured as the *Coefficient of Determination* (R²) that represents the proportion of variance explained. The model versions differed in their feature matrix: the base model used only the lesion map for prediction, whereas the other models used in addition the extent of white matter hyperintensities (WMH) as either ordinal visual scale ratings (*CHS* – Cardiovascular Health Study, *PV-WMH* – periventricular WMH, *DS-WMH* – deep subcortical WMH, *PV-WMH and DS-WMH* – both vectors concatenated) or voxel-based measures (volume or maps) as predictors. For more details on model versions, see Fig. 1B. For each model version, prediction accuracy is depicted for all 10 repetitions using different sample randomizations and for the averaged model ('Mean', predictions of model repetitions were averaged). Averaged models are highlighted with asterisks if permutation testing revealed that the original WMH information significantly improved model performance (** – $p < 0.01$; see Fig. 5).
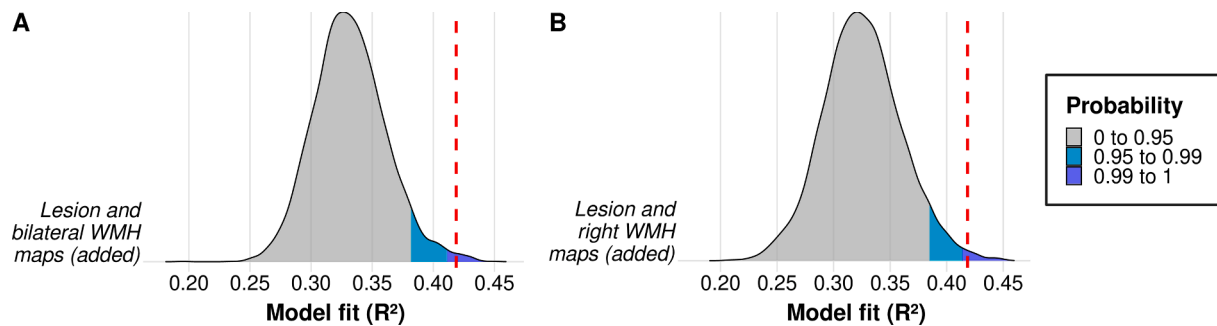
Fig. 5. Permutation tests revealing measures of WMH extent important for prediction. Distribution plots of 5,000 averaged models obtained by permutation testing are shown (predictions of permutation models were averaged across 10 model repetitions); WMH extent served as the permuted variable. Prediction accuracy was measured as the *Coefficient of Determination* (R^2) that represents the proportion of variance explained. R^2-distributions are displayed for the two model versions, for which the averaged model that used the true WMH data yielded a significant improvement of prediction accuracy above chance: addition of voxel-based volumetric maps containing (**A**) lesion and bilateral WMH and (**B**) lesion and right hemispheric WMH, indicating that bilateral and right hemispheric WMH maps were important features for prediction. The model fit of the corresponding original model is visualized as a red dashed line ($R^2 = 0.419$ or $R^2 = 0.418$, respectively). Areas of R^2-distribution that represent one-tailed significance are colored (grey: $p > 0.05$; cyan blue: $p < 0.05$; dark blue: $p < 0.01$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

WMH. However, permutation testing revealed no improvement of prediction accuracy above chance ($p > 0.05$).

3.2.2. Voxel-based WMH measures

In comparison to ordinal WMH ratings, the model that used WMH volumes resulted in a numerically similar prediction accuracy as the CHS and DS-WMH ratings, suggesting that WMH volume is no better predictor than ordinal ratings. In case of appending bilateral WMH maps, the prediction got more accurate, indicating that voxel-based WMH maps were more predictive than WMH volume and some rating-based WMH measures. However, the concatenation of lesion maps and bilateral WMH volumes or maps did not result in a numerical increase of prediction accuracy compared to the base model. Opposed to that, matrix addition of lesion map and bilateral WMH map yielded a numerically and also statistically significantly improved prediction accuracy of $R^2 = 0.419$ ($p = 0.007$; Fig. 5A), indicating that the matrix addition resulted in more accurate models than the matrix concatenation. When using unilateral volumetric WMH maps for matrix addition (Fig. S4A), no numerical improvement of prediction accuracy was found for the left hemispheric WMH map. In contrast, results demonstrated an improvement of performance significantly above chance for the right hemispheric WMH map ($R^2 = 0.418$; $p = 0.006$; Fig. 5B). The achieved model fit for the right hemispheric WMH map was about as large as for the bilateral WMH map (cf. Fig. 4 and Fig. 5), suggesting that right hemispheric WMH were more predictive than left hemispheric WMH and as sufficiently predictive as the whole WMH information. In sum, compared to the base model, the proportion of explained variance improved by approximately 5.1 % when adding the lesion map and the bilateral (or right hemispheric) WMH map together. This model was shown to be the overall best performing model, which explained nearly 42 % of the total variance.

3.2.3. Age

The observed improvement of including WMH (when adding lesion and bilateral or only right hemispheric WMH) in predicting spatial neglect might be simply an indirect effect of age, due to the known positive correlation between age and WMH extent. We thus asked whether a model using age instead of WMH data yields a similar prediction accuracy. For this purpose, we concatenated the lesion matrix and the vector of normalized age values like we did earlier for the WMH scale ratings. The averaged model that used lesion map and age as features revealed a prediction accuracy of $R^2 = 0.351$. Thus, appending age could explain little more variance than appending bilateral WMH volume and little less than appending the bilateral WMH map.

Therefore, appending the feature age also resulted in a less accurate model performance compared to the base model. Further, we calculated a model that attached age as a feature to the model with the best performance so far (i.e., lesion map and bilateral WMH map added). We yielded a model fit of $R^2 = 0.380$, which is again less accurate than the equivalent model without age as a feature. The results indicate that the WMH map is a better predictor than age.

4. Discussion

In the present study, we implemented predictive modeling using support vector regression (SVR) to clarify whether the extent of white matter hyperintensities (WMH) in addition to voxel-based stroke lesion information is predictive for the severity of spatial neglect in acute stroke patients. We compared different measures of WMH extent, namely rating-based measures assessed via ordinal visual rating scales and voxel-based (volumetric, topographic) measures quantified via a semi-automated demarcation procedure. The combination of volumetric lesion maps and WMH maps yielded improved accuracies compared to the stroke lesion volume-map alone. This finding supports observations by Bahrainwala and colleagues (2014) who observed an association between WMH severity and the probability of having any and also more severe spatial neglect after an acute right hemispheric stroke. The finding is also in line with another study demonstrating an improved accuracy in predicting cognitive deficits when including WMH extent compared to a base model without WMH information (Meier et al., 2020). We observed no increase in model performance when implementing age as an additional feature, although WMH extent and age were moderately correlated. Hence, our analyses revealed that this improvement in predicting neglect severity by including topographic WMH extent information is no indirect effect of age.

The severity of behavioral outcome and lesion size are correlated for many poststroke deficits (Sperber, 2022), and specifically for spatial neglect (Karnath et al., 2004). One could thus argue that the increase of accuracy in predicting spatial neglect by adding the WMH extent to the stroke lesion size is a trivial observation. However, the present findings demonstrate that this assumption does not hold. We observed that the prediction accuracy significantly improved when adding the voxel-based WMH map to the lesion map. In contrast, WMH volume did not achieve an improved model performance. Our findings hence indicate that the WMH volume per se did not generally benefit the model. Rather the combined information of volume and topography by means of a voxel-based map improved the prediction accuracy. The present results thus demonstrated that the subject-specific joint information of WMH

volume and location is essential for an improved prediction of neglect severity. Further, our findings suggest that the voxel-based WMH map is most relevant when treating it as part of the brain lesion itself via direct addition to the lesion map, and less (or even not at all) relevant when treating both maps separately. In other words, the anatomical information was most predictive when we treated WMH and the acute stroke lesion as if they constituted the same entity. This finding implies that WMH, although different in etiology, size, and chronicity compared to an acute major ischemic event, might be similar to a stroke lesion in regards to their impact on brain function.

Moreover, our study revealed that prediction via SVR using voxel-based WMH maps as topographic measures outperformed prediction using ordinal visual scale ratings that Bahrainwala and colleagues (2014) used in their study. A volumetrically quantified WMH map not only improved predictability by up to 1.7 % (when directly compared to scale ratings), but also provides a potentially more reliable measure. While our study detected WMH by a semi-automated algorithm, scale ratings – such as the CHS-scale (Manolio et al., 1994) or the Fazekas-scale (Fazekas et al., 1987) – are dependent on the visual evaluation of the individual rater and his or her interpretation of the specific scale levels. This circumstance is documented by the rather low consensus in terms of interrater reliability, especially for the Fazekas DS-WMH score in our study. A recent investigation that used both scales (CHS and Fazekas) reported comparable observations: interrater reliability was moderate for the CHS-scale and lower for the Fazekas-subscales (Meier et al., 2020). A further aspect arguing in favour of (semi-)automated algorithms for the detection of WMH might be that they allow to detect subtle WMH that would have been overlooked by purely visual inspection but might already constitute diffuse neuropathological changes (Grafton et al., 1991; Schmidt et al., 2011). Rating-based WMH measures, on the other hand, are not suited for detecting potentially relevant nuances and subtle patterns of brain damage.

4.1. Hemisphere effects

WMH of presumed vascular origin are mostly symmetrically in both hemispheres. Interestingly, our study revealed that only ipsilesional, right hemispheric (and bilateral) WMH maps, but not contralesional, left hemispheric WMH maps, were significantly important for the accuracy of predicting spatial neglect. In line with this view, Chen and colleagues (2020) reported that affected white matter fibers of especially the right hemisphere were linked to cognitive impairments in patients with WMH. What might our finding mean that right hemispheric WMH are more relevant than left hemispheric WMH? It is well known that the right hemisphere is dominant in spatial attention (e.g., Becker and Karnath, 2007) and damage to right hemispheric cortical and white matter areas underlie spatial neglect (Karnath et al., 2004; Karnath and Rorden, 2012; Vaessen et al., 2016; Wiesen et al., 2019). An alteration of the right hemispheric white matter microstructure that occurred already prior to stroke might lead to partial disconnections between neglect-related areas of the right hemisphere. The brain would henceforth be more vulnerable to develop attentional deficits poststroke as implied by the concept of 'brain reserve' (e.g., Umarova et al., 2021; for a review on brain reserve and spatial neglect, see Umarova, 2017). In a nutshell, the healthier and more sophisticated a brain is, the less it will be affected by events such as stroke, because the brain can tolerate more damage due to a greater reserve. The present study therefore contributes to the discussion on the impact of brain reserve on poststroke cognition and elucidates that WMH are critical for one of the most debilitating neuropsychological disorders after a right hemispheric stroke. On the other hand, it was previously shown that 'cognitive reserve' as measured by years of education is not predictive for spatial neglect (Umarova et al., 2019). In summary, the underlying mechanisms of these findings might contribute to the observation that severe ipsilesional, right hemispheric WMH are relevant for predicting deficits in visual attention after stroke.

4.2. Periventricular WMH versus deep subcortical WMH

Recent studies proposed the concept of 'white matter bottlenecks' whose associated areas appear to be critical in the development of cognitive impairment. These bottleneck regions correspond with the location of PV-WMH and were found to be predictive for language outcome poststroke (Griffis et al., 2017). This supports our conclusion that PV-WMH play a greater role than DS-WMH in predicting neglect severity. In line with this, previous research has reported that PV-WMH were related to deficits in cognition – as, e.g., visuospatial processing in Parkinson's disease (Huang et al., 2020) and language function in primary progressive aphasia (Meier et al., 2020). Compared to DS-WMH, a higher relative risk of having dementia was also attributed to PV-WMH (Georgakis et al., 2019). However, opposite results were obtained by Kamakura and colleagues (2017) who observed a stronger contribution of DS-WMH compared to PV-WMH in the relationship of severe WMH and worse subacute neglect.

Furthermore, we found only a weak correlation between PV-WMH and DS-WMH. This supports the previous assumption that both WMH variants might have separate origins (Armstrong et al., 2020; Wen et al., 2009) and argues against the suggestion of a common cause (DeCarli et al., 2005; Meier et al., 2020). We speculate that these discrepancies in the literature result not only from dissimilar cohorts but also from the various approaches to assessing WMH extent and the associated comparability issues as already mentioned in a *meta*-analysis (Armstrong et al., 2020); especially, visual rating scales causing disagreements in the interpretation of the severity levels and in the scoring of WMH extent (see above). In future research, the demarcation of WMH maps with (semi-)automated algorithms should have a priority.

4.3. WMH extent and chronic neglect

There is first evidence for the association of WMH extent and recovery from neglect. Hawe and colleagues (2018) reported a significant relation between chronic neglect performance and the calculated overall volumes of stroke lesion and WMH, whereas the volumes of lesion and of WMH alone were not associated with neglect recovery. Their finding corresponds well with our observation that the combination of stroke lesion and WMH volume-maps is a better predictor than one of those maps alone for neglect severity in acute stroke patients. Hawe and colleagues (2018) concluded that the stroke lesion alone did not cause chronic neglect as severe as when extensive WMH were combined with it. It is possible that a pre-damaged brain due to WMH formation might be more vulnerable to develop attentional deficits following stroke, maybe due to white matter disconnections. Further, Kamakura and colleagues (2017) found in a case series that less severe WMH were related to less severe neglect in the subacute phase; the degree of WMH extent differed between patients who had or had not recovered from neglect. In regard to a possible reorganization poststroke, the impact of WMH on the severity of spatial neglect is in line with the relevance of brain network damage found in previous studies not only for acute (Karnath et al., 2009; Wiesen et al., 2019) but also chronic spatial neglect (Lunven et al., 2015; Thiebaut de Schotten et al., 2014). Hence, WMH likely also affects the persistence of spatial neglect after stroke.

4.4. Limitations

Some limitations have to be considered. First, the present study used SVR models for prediction in accordance to previous works. Nevertheless, we cannot guarantee that SVR represents the best performing algorithm to predict spatial neglect. Future studies should investigate whether other (regression) models perform better than SVR and would therefore be more useful for future applications in clinical settings. Second, we set the threshold for keeping principal components during the PCA to 94 % cumulative explained variance since we found this threshold to work best for our sample. However, in other investigations,

a different cut-off might be optimal. For developing an algorithm suitable to predict real clinical data, it should further be considered to apply the PCA separately for training and test subsamples during the cross-validation procedure. Last, results and corresponding conclusions were found for the applied methological choices. Another cut-off for the PCA, for example, can yield slightly different outcomes and a deviating order of best performing WMH measures (see https://doi.org/10.1101/2022.04.08.22273547). Nevertheless, the effect of a significantly improved performance when adding WMH maps to the lesion maps appears to be stable.

4.5. Conclusions

The present study demonstrated that voxel-based WMH information affect the severity of spatial neglect in the acute phase of stroke. Our findings argue in favor of using combined maps of stroke lesion and WMH extent in predictive modeling. These maps might also lead to more accurate results in lesion-behavior mapping. Results further indicate that the combination of precise voxel-based maps of both stroke lesion and bilateral or ipsilesional, right hemispheric WMH yields an improved accuracy in predicting spatial neglect severity compared to the acute stroke lesion map alone. Since our models with precise volumetric WMH maps as auxiliary predictors resulted in better accuracies than ordinally scaled visual ratings (i.e., CHS-scale and the scale by Fazekas), we recommend measuring WMH extent topographically by means of a voxel-based map and combine it via direct matrix addition to the stroke lesion map. Further studies are needed that apply the approach of the present research to investigate if the voxel-based WMH map results also in a more accurate prediction of other cognitive deficits following stroke. If so, the methodological approach applied and associated results have great potential to become clinically relevant for rehabilitation as WMH appears to constitute a critical imaging biomarker for (poststroke) impairment in cognitive functions. Overall, the present findings might help to enhance the awareness of WMH impact and foster future (AI-based) algorithms to provide more accurate clinical diagnoses and predictions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data containing patient information cannot be made publicly available due to data protection restrictions by the local ethics commission. Remaining data including custom scripts are openly available in Mendeley Data (doi: 10.17632/c8n42jz525.1).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2022.103265.

References

Andersen, S.D., Larsen, T.B., Gorst-Rasmussen, A., Yavarian, Y., Lip, G.Y.H., Bach, F.W., 2017. White Matter Hyperintensities Improve Ischemic Stroke Recurrence Prediction. Cerebrovasc Dis. 43 (1–2), 17–24. https://doi.org/10.1159/000450962.

Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. Statistics Surveys. 4(none):40–79 https://doi.org/10.1214/09-SS054.

Armstrong, N.J., Mather, K.A., Sargurupremraj, M., et al., 2020. Common Genetic Variation Indicates Separate Causes for Periventricular and Deep White Matter Hyperintensities. Stroke. 51 (7), 2111–2121. https://doi.org/10.1161/STROKEAHA.119.027544.

Bahrainwala, Z.S., Hillis, A.E., Dearborn, J., Gottesman, R.F., 2014. Neglect performance in acute stroke is related to severity of white matter hyperintensities. Cerebrovasc Dis. 37 (3), 223–230. https://doi.org/10.1159/000357661.

Basilakos, A., Stark, B.C., Johnson, L., et al., 2019. Leukoaraiosis is associated with a decline in language abilities in chronic aphasia. Neurorehabil Neural Repair. 33 (9), 718–729. https://doi.org/10.1177/1545968319862561.

Becker, E., Karnath, H.O., 2007. Incidence of visual extinction after left versus right hemisphere stroke. Stroke. 38 (12), 3172–3174. https://doi.org/10.1161/STROKEAHA.107.489096.

Brereton, R.G., Lloyd, G.R., 2010. Support vector machines for classification and regression. Analyst. 135 (2), 230–267. https://doi.org/10.1039/b918972f.

Chang, C.C., Lin, C.J., 2011. LIBSVM: A library for support vector machines. ACM Trans Intell Syst Technol. 2(3):27: 1 –27, 27. https://doi.org/10.1145/1961189.1961199.

Chen, H.F., Huang, L.L., Li, H.Y., et al., 2020. Microstructural disruption of the right inferior fronto-occipital and inferior longitudinal fasciculus contributes to WMH-related cognitive impairment. CNS Neurosci Ther. 26 (5), 576–588. https://doi.org/10.1111/cns.13283.

Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput Sci. 7, e623.

Clas, P., Groeschel, S., Wilke, M., 2012. A semi-automatic algorithm for determining the demyelination load in metachromatic leukodystrophy. Acad Radiol. 19 (1), 26–34. https://doi.org/10.1016/j.acra.2011.09.008.

Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J Comput Assist Tomogr. 18 (2), 192–205.

de Haan, B., Clas, P., Juenger, H., Wilke, M., Karnath, H.O., 2015. Fast semi-automated lesion demarcation in stroke. Neuroimage Clin. 9, 69–74. https://doi.org/10.1016/j.nicl.2015.06.013.

Debette, S., Markus, H.S., 2010. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. BMJ. 341, c3666 https://doi.org/10.1136/bmj.c3666.

DeCarli, C., Fletcher, E., Ramey, V., Harvey, D., Jagust, W.J., 2005. Anatomical Mapping of White Matter Hyperintensities (WMH). Stroke. 36 (1), 50–55. https://doi.org/10.1161/01.STR.0000150668.58689.f2.

Fazekas, F., Chawluk, J., Alavi, A., Hurtig, H., Zimmerman, R., 1987. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. American Journal of Roentgenology. 149 (2), 351–356. https://doi.org/10.2214/ajr.149.2.351.

Gauthier, L., Dehaut, F., Joanette, Y., 1989. The Bells Test: A quantitative and qualitative test for visual neglect. International Journal of Clinical Neuropsychology. 11, 49–54.

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(12): e1298-e1308. 10.1212/WNL.0000000000007142.

Grafton, S.T., Sumi, S.M., Stimac, G.K., Alvord Jr, E.C., Shaw, C.M., Nochlin, D., 1991. Comparison of Postmortem Magnetic Resonance Imaging and Neuropathologic Findings in the Cerebral White Matter. Archives of Neurology. 48 (3), 293–298. https://doi.org/10.1001/archneur.1991.00530150061019.

Griffis, J.C., Nenert, R., Allendorfer, J.B., Szaflarski, J.P., 2017. Damage to white matter bottlenecks contributes to language impairments after left hemispheric stroke. NeuroImage: Clinical. 14, 552–565. https://doi.org/10.1016/j.nicl.2017.02.019.

Guttmann, C.R., Benson, R., Warfield, S.K., et al., 2000. White matter abnormalities in mobility-impaired older persons. Neurology. 54 (6), 1277–1283. https://doi.org/10.1212/wnl.54.6.1277.

Hawe, R.L., Findlater, S.E., Kenzie, J.M., Hill, M.D., Scott, S.H., Dukelow, S.P., 2018. Differential Impact of Acute Lesions Versus White Matter Hyperintensities on Stroke Recovery. J Am Heart Assoc. 7 (18), e009360.

Hope, T.M.H., Leff, A.P., Price, C.J., 2018. Predicting language outcomes after stroke: Is structural disconnection a useful predictor? Neuroimage Clin. 19, 22–29. https://doi.org/10.1016/j.nicl.2018.03.037.

Huang, X., Wen, M.C., Ng, S.Y.E., et al., 2020. Periventricular white matter hyperintensity burden and cognitive impairment in early Parkinson's disease. European Journal of Neurology. 27 (6), 959–966. https://doi.org/10.1111/ene.14192.

Kamakura, C.K., Ueno, Y., Sakai, Y., et al., 2017. White matter lesions and cognitive impairment may be related to recovery from unilateral spatial neglect after stroke. J Neurol Sci. 379, 241–246. https://doi.org/10.1016/j.jns.2017.06.021.

Karnath, H.O., Dieterich, M., 2006. Spatial neglect–a vestibular disorder? Brain. 129 (Pt 2), 293–305. https://doi.org/10.1093/brain/awh698.

Karnath HO, Sperber C, Wiesen D, de Haan B. Lesion-Behavior Mapping in Cognitive Neuroscience: A Practical Guide to Univariate and Multivariate Approaches. In: Pollmann S, ed. *Spatial Learning and Attention Guidance.* Vol 151. Neuromethods. Springer US; 2019:209-238. 10.1007/7657_2019_18.

Karnath, H.O., Rorden, C., 2012. The anatomy of spatial neglect. Neuropsychologia. 50 (6), 1010–1017. https://doi.org/10.1016/j.neuropsychologia.2011.06.027.

Karnath, H.O., Fruhmann Berger, M., Küker, W., Rorden, C., 2004. The anatomy of spatial neglect based on voxelwise statistical analysis: a study of 140 patients. Cereb Cortex. 14 (10), 1164–1172. https://doi.org/10.1093/cercor/bhh076.

Karnath, H.O., Rorden, C., Ticini, L.F., 2009. Damage to White Matter Fiber Tracts in Acute Spatial Neglect. Cerebral Cortex. 19 (10), 2331–2337. https://doi.org/10.1093/cercor/bhn250.

Kasties, V., Karnath, H.O., Sperber, C., 2021. Strategies for feature extraction from structural brain imaging in lesion-deficit modelling. Hum Brain Mapp. 42 (16), 5409–5422. https://doi.org/10.1002/hbm.25629.

Khan, M., Heiser, H., Bernicchi, N., et al., 2019. Leukoaraiosis Predicts Short-term Cognitive But not Motor Recovery in Ischemic Stroke Patients During Rehabilitation. J Stroke Cerebrovasc Dis. 28 (6), 1597–1603. https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.02.037.

Krstajic, D., Buturovic, L.J., Leahy, D.E., Thomas, S., 2014. Cross-validation pitfalls when selecting and assessing regression and classification models. Journal of Cheminformatics. 6 (1), 10. https://doi.org/10.1186/1758-2946-6-10.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics. 33 (1), 159–174.

Linortner, P., McDaniel, C., Shahid, M., et al., 2020. White Matter Hyperintensities Related to Parkinson's Disease Executive Function. Movement Disorders Clinical Practice. 7 (6), 629–638. https://doi.org/10.1002/mdc3.12956.

Liu, H., Deng, B., Xie, F., et al., 2021. The influence of white matter hyperintensity on cognitive impairment in Parkinson's disease. Ann Clin Transl Neurol. 8 (9), 1917–1934. https://doi.org/10.1002/acn3.51429.

Lunven, M., Thiebaut De Schotten, M., Bourlon, C., et al., 2015. White matter lesional predictors of chronic visual neglect: a longitudinal study. Brain. 138 (Pt 3), 746–760. https://doi.org/10.1093/brain/awu389.

Manolio, T.A., Kronmal, R.A., Burke, G.L., et al., 1994. Magnetic resonance abnormalities and cardiovascular disease in older adults. The Cardiovascular Health Study. *Stroke*. 25 (2), 318–327. https://doi.org/10.1161/01.str.25.2.318.

Meier, E.L., Breining, B.L., Sheppard, S.M., et al., 2020. White Matter Hyperintensities Contribute to Language Deficits in Primary Progressive Aphasia. Cogn Behav Neurol. 33 (3), 179–191. https://doi.org/10.1097/WNN.0000000000000237.

Odolil, A., Wright, A.E., Keator, L.M., et al., 2020. Leukoaraiosis Severity Predicts Rate of Decline in Primary Progressive Aphasia. Aphasiology. 34 (3), 365–375. https://doi.org/10.1080/02687038.2019.1594152.

Qi, X., Tang, H., Luo, Q., et al., 2019. White Matter Hyperintensities Predict Cognitive Decline: A Community-Based Study. Can J Neurol Sci. 46 (4), 383–388. https://doi.org/10.1017/cjn.2019.47.

Rondina, J.M., Filippone, M., Girolami, M., Ward, N.S., 2016. Decoding post-stroke motor function from structural brain imaging. NeuroImage: Clinical. 12, 372–380. https://doi.org/10.1016/j.nicl.2016.07.014.

Rorden, C., Brett, M., 2000. Stereotaxic display of brain lesions. Behavioural Neurology. 12 (4), 191–200.

Rorden, C., Karnath, H.O., 2010. A simple measure of neglect severity. Neuropsychologia. 48 (9), 2758–2763. https://doi.org/10.1016/j.neuropsychologia.2010.04.018.

Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., Karnath, H.O., 2012. Age-specific CT and MRI templates for spatial normalization. Neuroimage. 61 (4), 957–965. https://doi.org/10.1016/j.neuroimage.2012.03.020.

Salvalaggio, A., Grazia, D.F.D.M., Zorzi, M., Thiebaut de Schotten, M., Corbetta, M., 2020. Post-stroke deficit prediction from lesion and indirect structural and functional disconnection. Brain. 143 (7), 2173–2188. https://doi.org/10.1093/brain/awaa156.

Schmidt, R., Schmidt, H., Haybaeck, J., et al., 2011. Heterogeneity in age-related white matter changes. Acta Neuropathol. 122 (2), 171–185. https://doi.org/10.1007/s00401-011-0851-x.

Siegel JS, Ramsey LE, Snyder AZ, et al. Disruptions of network connectivity predict impairment in multiple behavioral domains after stroke. *Proceedings of the National Academy of Sciences*. 2016;113(30):E4367-E4376. 10.1073/pnas.1521083113.

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Statistics and Computing. 14 (3), 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88.

Sperber, C., 2022. The strange role of brain lesion size in cognitive neuropsychology. Cortex. 146, 216–226. https://doi.org/10.1016/j.cortex.2021.11.005.

Stemler, S.E., 2004. A Comparison of Consensus, Consistency, and Measurement Approaches to Estimating Interrater Reliability. Practical Assessment, Research, and Evaluation. 9 (4) https://doi.org/10.7275/96jp-xz07.

Tang, X., Jiang, L., Luo, Y., et al., 2021. Leukoaraiosis and acute ischemic stroke. Eur J Neurosci. 54 (6), 6202–6213. https://doi.org/10.1111/ejn.15406.

Thiebaut de Schotten, M., Tomaiuolo, F., Aiello, M., et al., 2014. Damage to White Matter Pathways in Subacute and Chronic Spatial Neglect: A Group Study and 2 Single-Case Studies with Complete Virtual "In Vivo" Tractography Dissection. Cerebral Cortex. 24 (3), 691–706. https://doi.org/10.1093/cercor/bhs351.

Umarova, R.M., 2017. Adapting the concepts of brain and cognitive reserve to post-stroke cognitive deficits: Implications for understanding neglect. Cortex. 97, 327–338. https://doi.org/10.1016/j.cortex.2016.12.006.

Umarova, R.M., Sperber, C., Kaller, C.P., et al., 2019. Cognitive reserve impacts on disability and cognitive deficits in acute stroke. J Neurol. 266 (10), 2495–2504. https://doi.org/10.1007/s00415-019-09442-6.

Umarova, R.M., Schumacher, L.V., Schmidt, C.S.M., et al., 2021. Interaction between cognitive reserve and age moderates effect of lesion load on stroke outcome. Sci Rep. 11 (1), 4478. https://doi.org/10.1038/s41598-021-83927-1.

Vaessen, M.J., Saj, A., Lovblad, K.O., Gschwind, M., Vuilleumier, P., 2016. Structural white-matter connections mediating distinct behavioral components of spatial neglect in right brain-damaged patients. Cortex. 77, 54–68. https://doi.org/10.1016/j.cortex.2015.12.008.

Valdés Hernández, M.D.C., Booth, T., Murray, C., et al., 2013. Brain white matter damage in aging and cognitive ability in youth and older age. Neurobiol Aging. 34 (12), 2740–2747. https://doi.org/10.1016/j.neurobiolaging.2013.05.032.

Varoquaux, G., Raamana, P.R., Engemann, D.A., Hoyos-Idrobo, A., Schwartz, Y., Thirion, B., 2017. Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. NeuroImage. 145, 166–179. https://doi.org/10.1016/j.neuroimage.2016.10.038.

Wardlaw, J.M., Smith, C., Dichgans, M., 2013a. Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging. Lancet Neurol. 12 (5), 483–497. https://doi.org/10.1016/S1474-4422(13)70060-7.

Wardlaw, J.M., Smith, E.E., Biessels, G.J., et al., 2013b. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. Lancet Neurol. 12 (8), 822–838. https://doi.org/10.1016/S1474-4422(13)70124-8.

Wardlaw, J.M., Valdés Hernández, M.C., Muñoz-Maniega, S., 2015. What are white matter hyperintensities made of? Relevance to vascular cognitive impairment. J Am Heart Assoc. 4 (6), 001140 https://doi.org/10.1161/JAHA.114.001140.

Wei, K., Tran, T., Chu, K., et al., 2019. White matter hypointensities and hyperintensities have equivalent correlations with age and CSF β-amyloid in the nondemented elderly. Brain Behav. 9 (12), e01457.

Weintraub, S., Mesulam, M.M., 1985. Mental state assessment of the young and elderly adults in behavioral neurology. In: Mesulam, M.M. (Ed.), Principles of Behavioral Neurology. FA Davis, Philadelphia, PA, pp. 71–123.

Wen, W., Sachdev, P.S., Li, J.J., Chen, X., Anstey, K.J., 2009. White matter hyperintensities in the forties: their prevalence and topography in an epidemiological sample aged 44–48. Hum Brain Mapp. 30 (4), 1155–1167. https://doi.org/10.1002/hbm.20586.

Wiesen, D., Sperber, C., Yourganov, G., Rorden, C., Karnath, H.O., 2019. Using machine learning-based lesion behavior mapping to identify anatomical networks of cognitive dysfunction: Spatial neglect and attention. Neuroimage. 201, 116000 https://doi.org/10.1016/j.neuroimage.2019.07.013.

Wilmskoetter, J., Marebwa, B., Basilakos, A., et al., 2019. Long-range fibre damage in small vessel brain disease affects aphasia severity. Brain. 142 (10), 3190–3201. https://doi.org/10.1093/brain/awz251.

Zhang, Y., Kimberg, D.Y., Coslett, H.B., Schwartz, M.F., Wang, Z., 2014. Multivariate lesion-symptom mapping using support vector regression. Hum Brain Mapp. 35 (12), 5861–5876. https://doi.org/10.1002/hbm.22590.