



Universiteit
Leiden
The Netherlands

In shape: a novel approach to white matter hyperintensity analysis

Kuhn, J.A.

Citation

Kuhn, J. A. (2024, November 21). *In shape: a novel approach to white matter hyperintensity analysis*. Retrieved from <https://hdl.handle.net/1887/4150233>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4150233>

Note: To cite this publication please use the final published version (if applicable).



CHAPTER

6

IDENTIFICATION OF DISTINCT BRAIN MRI PHENOTYPES AND THEIR ASSOCIATION WITH LONG-TERM DEMENTIA RISK IN COMMUNITY-DWELLING OLDER ADULTS

Jasmin Annica Keller, Sigurdur Sigurdsson, Barbara Schmitz Abecassis, Ilse M.J. Kant, Mark A. Van Buchem, Lenore J. Launer, Matthias Van Osch, Vilmundur Gudnason, Jeroen De Bresser.

Published in: *Neurology*. 2024;00:e209176.

<https://doi.org/10.1212/WNL.0000000000209176>

6.1 ABSTRACT

Individual brain MRI markers only show at best a modest association with long-term occurrence of dementia. Therefore, it is challenging to accurately identify individuals at increased risk for dementia. We aimed to identify different brain MRI phenotypes by hierarchical clustering analysis based on combined neurovascular and neurodegenerative brain MRI markers and to determine the long-term dementia risk within the brain MRI phenotype subgroups.

Hierarchical clustering analysis based on 32 combined neurovascular and neurodegenerative brain MRI markers in community-dwelling individuals of the Age-Gen/Environment Susceptibility Reykjavik Study was applied to identify brain MRI phenotypes. A Cox proportional hazards regression model was used to determine the long-term risk for dementia per subgroup.

We included 3056 participants and identified 15 subgroups with distinct brain MRI phenotypes. The phenotypes ranged from limited burden, mostly irregular white matter hyperintensity (WMH) shape and cerebral atrophy, mostly irregularly WMHs and microbleeds, mostly cortical infarcts and atrophy, mostly irregularly shaped WMH and cerebral atrophy to multi-burden subgroups. Each subgroup showed different long-term risks for dementia (min–max range hazard ratios (HRs) 1.01–6.18; mean time to follow-up 9.9 ± 2.6 years); especially the brain MRI phenotype with mainly WMHs and atrophy showed a large increased risk (HR 6.18, 95% CI 3.37–11.32).

Distinct brain MRI phenotypes can be identified in community-dwelling older adults. Our results indicate that distinct brain MRI phenotypes are related to varying long-term risks of developing dementia. Brain MRI phenotypes may in the future assist in an improved understanding of the structural correlates of dementia predisposition.

6.2 INTRODUCTION

Most older adults have brain changes on MRI, such as cerebral atrophy or manifestations of cerebral small vessel disease (SVD).¹ These brain MRI markers mostly represent late resulting damage of different underlying pathologies and are therefore largely unspecific. This makes differentiation of underlying pathology based on brain MRI challenging.¹ Moreover, some brain abnormalities detected on MRI are regarded as related to normal ageing. It is currently unknown if specific brain MRI phenotypes represent an increased risk for dementia. The ability to determine an individual's risk for dementia based on MRI may in the future be useful to determine patient prognosis and may aid in patient selection for future treatment studies. Common brain MRI markers of neurovascular and neurodegenerative diseases are white matter hyperintensities (WMHs), lacunes, microbleeds, enlarged perivascular spaces, and cerebral atrophy. These brain MRI markers have been studied previously and are associated with the occurrence of dementia.²⁻⁶ However, individual brain MRI markers only show at best a modest association with long-term occurrence of dementia.⁴ It, therefore, remains challenging to identify individuals who are at increased risk to develop dementia.² Because of heterogenous etiology and mixed pathologies, methods combining different brain MRI markers into one model may likely aid in a more detailed characterization of, potential prognostically relevant, so-called brain MRI phenotypes. In a previous study within our group (in a different cohort), we aimed to detect an increased stroke and mortality risk in patients with manifest arterial disease and analyzed brain MRI markers in a combined way using a hierarchical clustering approach, resulting in the identification of different brain MRI phenotypes.⁷ In that study, distinct brain MRI phenotypes were detected that were associated with a different risk of future stroke and mortality. These brain MRI phenotypes can aid to identify the structural correlates of predisposition to different disease outcome. The association of distinct brain MRI phenotypes with long-term dementia risk remains unknown. We therefore aimed to identify different brain MRI phenotypes in community-dwelling individuals by combined hierarchical clustering analysis based on neurovascular and neurodegenerative brain MRI markers. Within each of these brain MRI phenotype subgroups, we determined the long-term dementia risk.

6.3 METHODS

6.3.1 Participants and study design

The data set used for the current analysis was acquired as part of the population based Age-Gene/Environment Susceptibility (AGES) Reykjavik Study.⁸ The cohort study was originally established in 1967 to prospectively study cardiovascular disease

in a random sample from the general population in Iceland. Participants were born between 1907 and 1935 and were living in Reykjavik in 1967. Remaining participants of the cohort were randomly selected for a follow-up and underwent a baseline brain MRI scan between 2002 and 2006.

Baseline diagnosis of dementia was assessed in a 3-step process, as described previously.⁹ In short, participants underwent the Mini-Mental State Examination and the Digit Symbol Substitution Test. Participants were administered a second battery of diagnostic tests based on positive results in the previous tests and possibly a third stage, which included neurologic tests and a proxy interview.⁹ Based on these tests, participants were considered to have normal cognition, mild cognitive impairment, or dementia at baseline. Dementia diagnosis based on the Diagnostic and Statistical Manual, Fourth Edition, guidelines was made in a consensus meeting with a geriatrician, neurologist, neuropsychologist, and neuroradiologist.

Education level and smoking status were collected using questionnaires. The highest completed education level (primary school, secondary school, college, and university) was entered. Participants who never smoked were categorized as non-smokers, participants who smoked regularly and at least 100 cigarettes or 20 cigars in a lifetime were categorized as former smokers, and participants who currently smoke were categorized as current smokers. Height (in centimeters) and weight (in kilograms) were measured and used to calculate body mass index. Hypertension was based on self-report, use of antihypertensive medication, or based on the measurements of systolic blood pressure >140 mm Hg and/or diastolic blood pressure >90 mm Hg. A standard mercury sphygmomanometer was used to measure systolic and diastolic blood pressure; the mean of 2 measurements was calculated. Diabetes mellitus was based on self-report of diabetes, use of antidiabetic medication, or fasting blood glucose level >7.0 mmol/L. Coronary artery disease was based on self-report plus the use of nitrates or evidence of a myocardial infarction on electrocardiogram.

Participants were followed from the date of the baseline MRI scan until diagnosis of dementia, loss to follow-up, or end of follow-up. Loss to follow-up means that the participants died or could not be contacted. Tracking for dementia diagnosis was done through vital statistics and hospital records and by the nursing home and home-based resident assessment instrument. The dementia follow-up of the AGES Reykjavik Study was concluded in 2015 (end of follow-up). The inclusion and exclusion of participants from the AGES Reykjavik Study for the current study is illustrated in Figure 6.1. For example, participants who were demented at baseline were excluded.

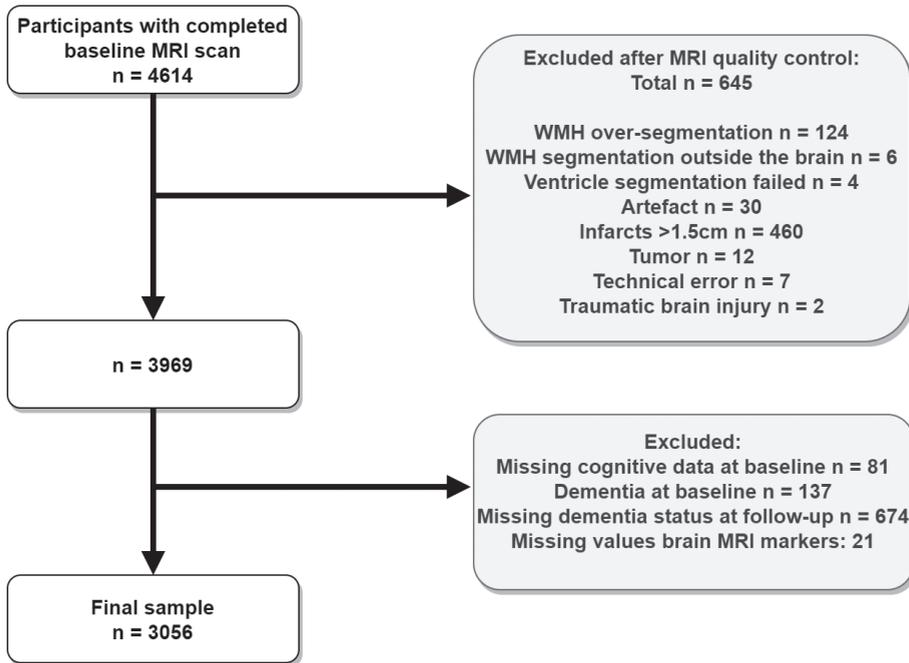


Figure 6.1 Flowchart Illustrating the Inclusion and Exclusion of Participants.

6.3.2 Standard protocol approvals, registrations, and patient consents

The study was approved by the Icelandic National Bioethics Committee, VSN:00-063, and the institutional review board responsible for the National Institute on Aging research; all participants signed for informed consent.

6.3.3 MRI Scanning protocol

A baseline brain MRI scan was acquired on a 1.5 T Signa TwinSpeed system (General Electric Medical Systems, Waukesha, WI). The MRI protocol included a fluid-attenuated inversion recovery (FLAIR) sequence (repetition time = 8000 milliseconds; time to echo = 100 milliseconds; inversion time = 2000 milliseconds; field of view = 220 mm; voxel size = $0.86 \times 0.86 \times 3.00 \text{ mm}^3$; interleaved slices) and a T1-weighted sequence (repetition time = 21 milliseconds; time to echo = 8 milliseconds; field of view = 240 mm; slice thickness = 1.5 mm; voxel size = $0.94 \times 0.94 \times 1.50 \text{ mm}^3$).¹¹ A T2*-weighted gradient echo-type echo planar sequence (time to echo = 50 milliseconds; repetition time = 3050 milliseconds; flip angle = 90° ; field of view = 220 mm; matrix = 256×256) and a proton density/T2-weighted fast-spin echo sequence (time to first echo = 22 milliseconds; time to second echo = 90 milliseconds; repetition time = 3220 milliseconds; echo train length = 8; flip angle = 90° ; field of view = 220 mm; matrix = 256×256) were also part of the MRI protocol.

6.3.4 Brain MRI markers

The brain MRI markers included to determine the brain MRI phenotypes were brain tissue volumes for the estimation of brain atrophy, WMH volumes, WMH shape markers, brain infarcts, microbleeds, and enlarged perivascular spaces. Gray matter, white matter, CSF, and WMH were segmented automatically with a modified algorithm based on the Montreal Neurological Institute pipeline.¹² Intracranial volume was calculated by adding gray matter, white matter, CSF, and WMH volumes.¹³ Brain parenchymal fraction, white matter fraction, gray matter fraction, and lateral ventricle fraction were calculated by expressing the volumes as a fraction of intracranial volume.

Volumes of periventricular/confluent, deep, and total WMH were determined automatically using an in-house developed pipeline.¹⁴ Moreover, volumes of deep and periventricular WMH per lobe were calculated using a mask to delineate the lobes. WMH shape markers (fractal dimension, solidity, convexity, concavity index, and eccentricity¹⁵) were calculated, as previously described.¹⁴ A description of the shape markers, as well as their corresponding formulas, can be found in supplementary table S.6.8.1. and supplementary figure S.6.8.1. (links.lww.com/WNL/D459). Brain infarcts (subcortical, cerebellar, and cortical infarcts), microbleeds, and enlarged perivascular spaces were visually scored.⁹ Microbleeds were first scored by neuroradiologists and then by trained radiographers.¹¹ Infarcts were defined as parenchymal defects with a signal intensity that is isointense to that of CSF on all MRI sequences (i.e., FLAIR, T2-weighted, proton density-weighted).⁹ Cortical infarcts were defined as infarcts involving or limited to the cortical gray matter and surrounded by a high signal intensity area on FLAIR images. Subcortical infarcts were categorized as such when they do not extend into the cortex and are surrounded by a high signal intensity area on FLAIR images of ≥ 4 mm in diameter.

Parenchymal defects in the subcortical area without a rim or area of high signal intensity on FLAIR images and without evidence of hemosiderin on the T2*-weighted scan were scored as enlarged perivascular spaces. Enlarged perivascular spaces were excluded from the definition of subcortical infarcts. Enlarged perivascular spaces were documented separately in the whole brain and in the basal ganglia. Cerebellar infarcts were scored without any size criteria. Infarcts covering 2 of the mentioned areas were attributed to the location in which the largest measured diameter was located (in millimeters).⁹

6.3.5 Statistical analysis

6.3.5.1 Identification of subgroups with different brain MRI phenotypes

All brain MRI markers were normalized as z-scores (after multiplication by 100 and natural log transformation when not normally distributed). Variables that were not normally distributed were WMH volumes, solidity, number of WMH, and lateral ventricle volume fraction. Binary variables (presence of microbleeds, infarcts, and enlarged perivascular spaces) were used as -2 and 2 to approximate the z-score distributions of continuous variables. Hierarchical clustering was performed by applying Ward's method in R version 4.1.0 (R Core Team, 2021) and packages *factoextra*,¹⁶ *cluster*,¹⁷ and *dendextend*¹⁸ on 32 brain MRI markers. Hierarchical clustering groups participants together based on similarities in brain MRI markers. The approach starts with every participant as a separate cluster and then repeatedly merging of the 2 closest clusters, subsequently updating the distance matrix. Thus, each cluster is the result of the merge of 2 subclusters, resulting in a hierarchical tree (dendrogram, Figure 6.2). At each level of the dendrogram, clusters are joined and the number of clusters therefore decreases. This is repeated until only 1 cluster, representing the total group of participants, remains. An optimal number of clusters need to be determined for further analysis. In an optimally clustered data set, the clusters have a high within-cluster cohesion, while having a high separation between different clusters. The optimal dendrogram cutoff, that is, the optimal number of clusters, was determined using the Dunn index (supplementary figure S.6.8.2., [links.lww.com/WNL/D459](https://www.lww.com/WNL/D459)) and the heatmap (Figure 6.2). The Dunn index is the ratio of the smallest distance between observations in different clusters over the largest between cluster distance and should be maximal. After this procedure, a number of subgroups remained, representing the subgroups with different brain MRI phenotypes.

Brain MRI markers and cardiovascular risk factors were compared between subgroups with chi-squared test for binary variables and 1-way analysis of variances for continuous variables by using SPSS version 25 (Chicago, IL). For these analyses, WMH volumes, number of WMH, solidity, ventricle volume fraction, and time to follow-up were log transformed due to a non-normal distribution. A p value <0.05 was considered statistically significant.

6.3.5.2 Sensitivity analysis

To assess the robustness of the hierarchical clustering model, we reran the analysis with 2 random subsets of this dataset.

6.3.5.3 Long-term outcome assessment

A Cox proportional hazard model was used to estimate the risk of future dementia occurrence within the brain MRI phenotype subgroups (adjusted for age, sex, and cognitive status at baseline (mild cognitive impairment or normal cognition). The reference subgroup was chosen based on having the fewest brain abnormalities. SPSS version 25 was used for statistical testing.

Table 6.1. Baseline characteristics of the total study sample.

Community-dwelling individuals (n = 3056)	
Age (years)	75.6 ± 5.2
Time to follow-up (years)	9.9 ± 2.6
Female sex	1884 (62%)
Mild cognitive impairment	257 (8%)
BMI (kg/m ²)	26.97 ± 4.19
Hypertension	2380 (78%)
Type 2 diabetes	277 (9%)
Cholesterol (mmol/L)	5.71 ± 1.14
Smoking status	
Never	1397 (46%)
Former	1334 (44%)
Current	324 (11%)
Coronary artery disease	497 (16%)

Data are shown as means ± SD or percentages of individuals per subgroup. BMI, body mass index; SD, standard deviation.

6.3.6 Data availability

The AGES I-II data set cannot be made publicly available because the informed consent signed by the participants prohibits data sharing on an individual level, as outlined by the study approval by the Icelandic National Bioethics Committee. Requests for these data may be sent to the AGES Reykjavik Study Executive Committee, contact: Ms. Camilla Kritjansdottir, Camilla@hjarta.is.

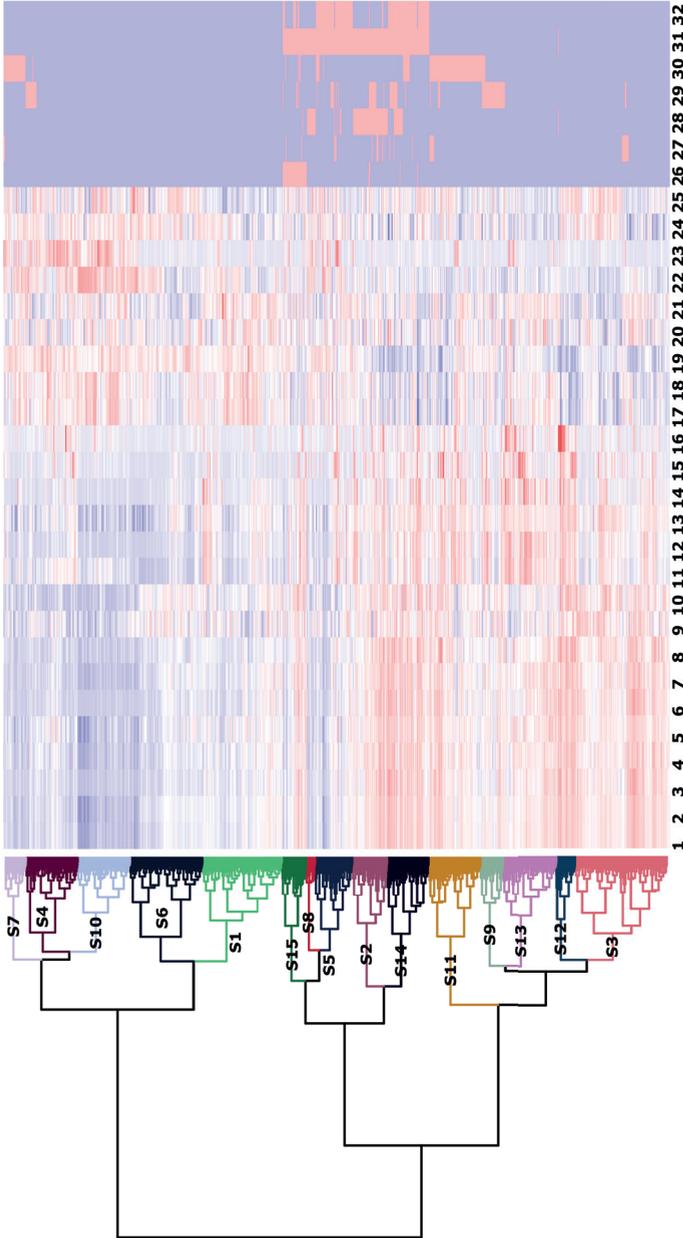


Figure 6.2 Heatmap and Dendrogram: Results of the Hierarchical Clustering Analysis

Red indicates high values, and blue indicates low values. Every line represents a participant, and every column represents a MRI marker. 1: total WMH volume; 2: PV/C WMH volume; 3: PV/C WMH volume frontal lobe; 4: other PV/C WMH volume; 5: PV/C WMH fractal dimension; 6: PV/C WMH volume parietal lobe; 7: PV/CWMH volume temporal lobe; 8: PV/C WMH concavity index; 9: % lateral ventricle volume; 10: PV/C WMH volume occipital lobe; 11: deep WMH volume; 12: deep WMH volume frontal lobe; 13: number of deep WMH; 14: deep WMH volume parietal lobe; 15: other deep WMH volume; 16: deep WMH volume temporal lobe; 17: %total brain volume; 18: % gray matter volume; 19: PV/C convexity; 20: % white matter volume; 21: deep WMH eccentricity; 22: deep WMH fractal dimension; 23: PV/C solidity; 24: deep WMH occipital lobe; 25: number of PV/C; 26: enlarged perivascular spaces in and around basal ganglia; 27: microbleeds; 28: cortical infarcts; 29: subcortical infarcts; 30: enlarged perivascular spaces in the white matter; 31: infarcts (whole brain); 32: cerebellar infarcts. Other PV/C WMH volume or other deep WMH volume were defined as Lesions or parts of lesions that were outside of the brain lobe masks, for example the brain stem and within the internal capsule. PV/C = periventricular/confluent; S = subgroup; WMH = white matter hyperintensity.

6.4 RESULTS

The total sample included 3056 community-dwelling older adults. The average time to follow-up for dementia outcome (yes/ no) was 9.9 ± 2.6 years (min–max range 0.6–13.4 years). The baseline characteristics of the total sample are presented in Table 6.1. A hierarchical clustering model was applied on brain MRI markers (WMH volumes, brain volumes, WMH shape markers, infarcts, enlarged perivascular spaces, microbleeds). The optimal cut-off of the hierarchical clustering model was determined to be at 15 subgroups (Figure 6.2), based on the Dunn index (supplementary figure S.6.8.2., links.lww.com/WNL/D459) and the heatmap (Figure 6.2). The sizes of the subgroups ranged from 42 to 425 participants. As a sensitivity analysis to test the robustness of the clustering method, we reran the model on 2 random subsets (subset 1: $n = 2,311$ and subset 2: $n = 2,250$). On average, 68% of participants remained in the same cluster compared with the main analysis. Baseline characteristics of the study sample per subgroup are shown in Table 2. The subgroups differed significantly in age (min–max range of mean age 71.6–78.8 years), sex (min–max range of mean sex distribution 37%–81% females), cognitive status at baseline (min–max range of mean prevalence 3%–15%), hypertension (min–max range of mean prevalence 64%–90%), type 2 diabetes mellitus (min–max range of mean prevalence 4%–15%), cholesterol levels (min–max range of mean 5.20–5.97 mmol/L), coronary artery disease (min–max range of mean prevalence 7%–30%), and time to follow-up (min–max range of mean time to follow-up 8.2–10.9 years; Table 6.2).

Brain MRI markers per subgroup are presented in Table 6.3. All brain MRI markers differed significantly between subgroups, as could be expected as the subgroups were based on the hierarchical clustering result. The main MRI markers per subgroup are illustrated in a simplified and summarized manner in Figure 6.3. The brain MRI phenotypes of the subgroups ranged from limited burden (subgroup 10), mostly irregular WMH shape and cerebral atrophy (subgroup 12), mostly irregularly shaped WMH and microbleeds (subgroup 9), mostly cortical infarcts and atrophy (subgroup 15), mostly irregularly shaped WMH and cerebral atrophy (subgroup 3) to multi-burden subgroups (subgroup 2, subgroup 14). A complete and detailed description of the main MRI markers of each subgroup can be found in S.6.8.1 supplementary results (links.lww.com/WNL/D459).

Subgroup 10 was determined to have the least amount of brain abnormalities and was used as the reference subgroup in the survival analysis (Figure 6.4). Dementia cases at follow-up ranged from 6% to 46% per subgroup. Compared with the reference subgroup, most other subgroups (except subgroup 4, subgroup 7, and subgroup 8) showed a higher long-term risk for dementia. The range of statistically

significant hazard ratios (HRs) across the subgroups varied between 1.01 and 6.18. The subgroup with relatively severe WMH and atrophy (subgroup 12) showed the largest long-term risk for dementia (HR 6.18, 95% CI 3.37–11.32) compared with the reference subgroup (subgroup 10). The multi-burden subgroups (subgroup 2: HR 3.68, 2.06–6.61; subgroup 14: HR 3.48, 1.96–6.16), the subgroup with mostly irregularly shaped WMH and cerebral atrophy (subgroup 3: HR 3.33, 1.94–5.73), the subgroup with mostly irregularly shaped WMH and microbleeds (subgroup 9: HR 2.99, 1.60–5.56), and the subgroup with mostly cortical infarcts and atrophy (subgroup 15: HR 3.16, 1.70–5.86) also showed a relatively high long-term risk for dementia.

	WMH volumes	WMH shape	Cerebellar infarcts	Cortical infarcts	Subcortical infarcts	Atrophy	Microbleeds	Enlarged PVS
S1	○	⊙	●	●	●	○	●	●
S2	○	⊙	●	●	●	○	●	●
S3	○	⊙	●	●	●	○	●	●
S4	○	○	●	●	●	○	●	●
S5	○	⊙	●	●	●	○	●	●
S6	○	⊙	●	●	●	○	●	●
S7	○	○	●	●	●	○	●	●
S8	○	○	●	●	●	○	●	●
S9	○	⊙	●	●	●	○	●	●
S10	○	○	●	●	●	○	●	●
S11	○	⊙	●	●	●	○	●	●
S12	○	⊙	●	●	●	○	●	●
S13	○	⊙	●	●	●	○	●	●
S14	○	⊙	●	●	●	○	●	●
S15	○	⊙	●	●	●	○	●	●

Figure 6.3. Symbolic illustration of the brain MRI markers per subgroup. Small circles indicate low WMH volumes or low amount of cerebral atrophy. Big circles indicate high WMH volumes or cerebral atrophy. WMH shape is illustrated with symbols with three varying degrees of shape irregularity (regular, moderate, and irregular). Pie charts indicate percentages of participants with infarcts, enlarged PVS or microbleeds. S: subgroup; WMH: white matter hyperintensities; enlarged PVS: enlarged perivascular spaces in and around the basal ganglia.

Table 6.2. Baseline characteristics of the study sample per subgroup. S: subgroup; BMI: body mass index.

	S1 N=368	S2 N=161	S3 N=425	S4 N=239	S5 N=171	S6 N=333	S7 N=98
Age (years)	74.4 ± 4.7	77.5 ± 5.3	77.8 ± 4.9	73.4 ± 4.1	75.6 ± 5.1	74.3 ± 4.8	74.2 ± 4.8
Time to dementia follow-up (years)	10.33 ± 2.02	8.89 ± 2.80	9.27 ± 2.93	10.89 ± 1.83	9.94 ± 2.33	10.36 ± 2.25	10.41 ± 2.33
Female sex	73%	52%	50%	74%	65%	54%	70%
Mild cognitive impairment	7%	11%	12%	3%	9%	5%	3%
BMI (kg/m ²)	26.74 ± 4.29	27.35 ± 4.29	27.02 ± 4.47	27.18 ± 4.07	26.85 ± 4.11	26.53 ± 4.13	26.99 ± 4.07
Hypertension	75%	88%	81%	76%	80%	70%	68%
Type 2 diabetes	6%	14%	10%	6%	10%	6%	9%
Cholesterol (mmol/L)	5.82 ± 1.10	5.44 ± 1.08	5.67 ± 1.15	5.91 ± 1.16	5.55 ± 1.20	5.77 ± 1.04	5.75 ± 1.10
Smoking status							
Never	49%	35%	45%	47%	44%	44%	54%
Former	41%	53%	47%	42%	44%	44%	38%
Current	9%	11%	8%	10%	12%	12%	8%
Coronary artery disease	10%	24%	16%	14%	19%	10%	10%

Data are shown as means ± SD or percentages of individuals per subgroup. Baseline characteristics and cardiovascular risk factors were compared between subgroups with chi-square test for binary variables, and one-way ANOVAs for continuous variables. S, subgroup; BMI, body mass index; SD, standard deviation.

S8 N=42	S9 N=105	S10 N=241	S11 N=240	S12 N=87	S13 N=245	S14 N=190	S15 N=111	P value
72.7 ± 3.2	76.8 ± 4.8	71.6 ± 3.4	76.5 ± 5.0	78.2 ± 4.9	75.6 ± 5.0	78.8 ± 5.3	77.3 ± 5.1	<0.001
10.53 ± 1.66	9.43 ± 2.68	10.70 ± 1.54	9.54 ± 2.89	8.18 ± 3.10	10.28 ± 2.50	8.78 ± 2.85	9.14 ± 2.77	<0.001
62% 5%	54% 12%	81% 5%	61% 9%	37% 11%	79% 8%	45% 15%	43% 15%	<0.001 <0.001
27.18 ± 4.28	25.91 ± 3.52	27.18 ± 3.99	26.90 ± 4.19	28.41 ± 4.52	27.29 ± 4.55	27.12 ± 3.51	26.72 ± 3.93	0.021
79% 17%	84% 10%	64% 4%	83% 9%	90% 11%	82% 11%	88% 15%	82% 15%	<0.001 <0.001
5.63 ± 0.87	5.69 ± 1.22	5.97 ± 1.07	5.76 ± 1.23	5.52 ± 1.08	5.75 ± 1.14	5.48 ± 1.22	5.20 ± 1.12	0.023
36%	45%	54%	49%	38%	43%	44%	44%	0.034
45%	44%	39%	40%	47%	42%	46%	45%	0.295
19%	11%	7%	11%	15%	15%	10%	11%	0.214
7%	21%	7%	21%	23%	18%	29%	30%	<0.001

Table 6.3. Between-group differences of brain MRI markers.

	S1 N=368	S2 N=161	S3 N=425	S4 N=239	S5 N=171	S6 N=333	S7 N=98
Total WMH volume	11.87 ± 4.69	35.85± 22.86	29.93 ± 16.76	7.50 ± 3.33	11.76 ± 7.66	8.02 ± 3.72	5.77 ± 2.25
PV/C WMH volume	10.68 ± 4.59	33.76 ± 22.83	28.51 ± 16.81	6.05 ± 3.10	10.28 ± 7.04	7.57 ± 3.60	4.83 ± 2.14
Deep WMH volume	1.18 ± 0.94	2.09 ± 1.52	1.42 ± 0.91	1.45 ± 0.85	1.49 ± 1.69	0.45 ± 0.39	0.94 ± 0.81
Number of PV/C	14.69 ± 5.70	14.86± 5.03	14.96 ± 5.51	14.90 ± 5.89	14.71 ± 5.07	15.27 ± 6.19	15.53 ± 6.06
Number of deep WMH	25.40 ± 15.29	44.69 ± 29.64	41.26 ± 25.69	21.23 ± 13.84	28.89 ± 26.73	16.39 ± 13.72	15.06 ± 10.83
PV/C solidity	0.18 ± 0.07	0.15 ± 0.05	0.13 ± 0.05	0.23 ± 0.11	0.21 ± 0.13	0.20 ± 0.11	0.25 ± 0.11
PV/C convexity	1.11 ± 0.12	0.87 ± 0.18	0.89 ± 0.14	1.13 ± 0.13	1.08 ± 0.14	1.13 ± 0.12	1.13 ± 0.11
PV/C concavity index	1.22 ± 0.07	1.42 ± 0.14	1.41 ± 0.11	1.17 ± 0.07	1.22 ± 0.11	1.19 ± 0.08	1.15 ± 0.07
PV/C fractal dimension	1.69 ± 0.09	1.84 ± 0.11	1.82 ± 0.08	1.65 ± 0.11	1.67 ± 0.14	1.62 ± 0.11	1.58 ± 0.12
Deep WMH eccentricity	0.56 ± 0.06	0.61 ± 0.06	0.63 ± 0.06	0.61 ± 0.07	0.62 ± 0.08	0.65 ± 0.08	0.62 ± 0.08
Deep WMH fractal dimension	1.79 ± 1.79	1.69 ± 0.12	1.65 ± 0.12	1.71 ± 0.13	1.69 ± 0.13	1.59 ± 0.12	1.69 ± 0.15
% lateral ventricle volume	2.61 ± 0.97	3.23 ± 1.09	3.60 ± 1.15	2.18 ± 0.72	2.63 ± 1.08	2.82 ± 0.93	2.19 ± 0.72
% total brain volume	74.52 ± 3.02	71.48 ± 3.59	70.80 ± 3.02	74.27 ± 2.45	72.91 ± 3.07	72.25 ± 2.82	73.44 ± 3.62
% grey matter volume	46.84 ± 2.60	44.50 ± 2.96	43.97 ± 2.62	47.28 ± 2.04	45.95 ± 2.70	45.41 ± 2.43	46.49 ± 3.09
% white matter volume	27.68 ± 1.49	26.98 ± 1.93	26.83 ± 1.35	26.98 ± 1.55	26.96 ± 1.47	26.84 ± 1.49	26.95 ± 1.54
PVS (basal ganglia)	0%	9%	8%	0%	5%	0%	4%
PVS	0%	0%	0%	3%	10%	0%	100%
Microbleeds	0%	14%	1%	16%	12%	0%	0%
Subcortical infarcts	0%	100%	0%	0%	8%	0%	0%
Cerebellar infarcts	0%	5%	0%	0%	98%	0%	0%
Cortical infarcts	0%	4%	0%	0%	1%	0%	0%
Infarcts	0%	100%	0%	0%	100%	0%	0%

Data shown as mean ±SD or percentages of individuals per subgroup. Brain MRI markers were compared between subgroups with chi-square Test for binary variables, and one-way ANOVAs for continuous variables. The data for WMH volumes per lobe are shown in supplementary table S.6.8.2. S, subgroup; WMH, white matter hyperintensity; PV/C periventricular/confluent WMH; PVS, enlarged perivascular spaces; SD, standard deviation.

S8 N=42	S9 N=105	S10 N=241	S11 N=240	S12 N=87	S13 N=245	S14 N=190	S15 N=111	P value
6.07 ± 2.66	22.90 ± 12.89	3.20 ± 1.44	25.93 ± 15.89	50.32 ± 20.71	24.56 ± 9.63	41.11 ± 21.16	22.85 ± 16.85	<0.001
5.18 ± 2.65	21.05 ± 12.86	2.37 ± 1.29	23.93 ± 15.72	46.99 ± 20.77	21.43 ± 9.71	39.19 ± 21.46	21.29 ± 16.60	<0.001
0.89 ± 0.75	1.85 ± 1.41	0.83 ± 0.51	2.00 ± 1.27	3.33 ± 1.51	3.13 ± 1.81	1.91 ± 1.36	1.57 ± 1.35	<0.001
15.36 ± 5.94	14.78 ± 5.57	16.00 ± 5.98	14.93 ± 6.08	15.62 ± 6.12	15.23 ± 5.94	14.78 ± 5.46	14.57 ± 5.41	<0.001
14.67 ± 10.29	41.90 ± 22.40	8.20 ± 5.11	46.33 ± 26.79	86.90 ± 40.35	45.49 ± 16.62	46.88 ± 28.29	34.59 ± 23.67	<0.001
0.25 ± 0.12	0.13 ± 0.05	0.39 ± 0.16	0.14 ± 0.05	0.03 ± 0.15	0.14 ± 0.04	0.14 ± 0.05	0.17 ± 0.07	<0.001
1.15 ± 0.11	0.97 ± 0.16	1.12 ± 0.11	0.92 ± 0.19	0.67 ± 0.13	0.97 ± 0.14	0.81 ± 0.16	1.00 ± 0.17	<0.001
1.15 ± 0.06	1.35 ± 0.12	1.09 ± 0.06	1.38 ± 0.15	1.59 ± 0.11	1.34 ± 0.10	1.47 ± 0.13	1.31 ± 0.14	<0.001
1.58 ± 0.11	1.78 ± 0.10	1.48 ± 0.14	1.80 ± 0.11	1.92 ± 0.07	1.81 ± 0.09	1.87 ± 0.09	1.74 ± 0.12	<0.001
0.60 ± 0.10	0.61 ± 0.07	0.64 ± 0.08	0.61 ± 0.06	0.66 ± 0.05	0.58 ± 0.05	0.63 ± 0.06	0.62 ± 0.06	<0.001
1.68 ± 0.12	1.68 ± 0.13	1.72 ± 0.13	1.69 ± 0.13	1.60 ± 0.11	1.77 ± 0.08	1.66 ± 0.12	1.70 ± 0.12	<0.001
2.21 ± 0.91	3.09 ± 1.21	1.85 ± 0.64	2.92 ± 1.06	3.43 ± 0.98	2.50 ± 0.82	3.64 ± 1.21	3.08 ± 1.07	<0.001
74.93 ± 3.16	72.02 ± 3.56	75.73 ± 3.07	71.94 ± 3.60	68.76 ± 4.15	74.55 ± 2.93	70.69 ± 4.01	71.75 ± 3.09	<0.001
47.86 ± 2.32	45.06 ± 2.85	48.20 ± 2.58	44.96 ± 2.93	42.71 ± 3.40	46.76 ± 2.41	43.72 ± 3.38	44.45 ± 2.83	<0.001
27.07 ± 1.35	26.96 ± 1.74	27.53 ± 1.48	26.98 ± 1.70	26.05 ± 2.03	27.79 ± 1.63	26.97 ± 1.85	27.30 ± 1.59	<0.001
0%	0%	0%	8%	2%	0%	6%	8%	<0.001
0%	17%	0%	99%	2%	0%	17%	16%	<0.001
5%	66%	0%	2%	1%	0%	9%	7%	<0.001
100%	0%	0%	0%	3%	0%	22%	5%	<0.001
0%	0%	0%	0%	0%	0%	99%	29%	<0.001
0%	0%	0%	0%	0%	0%	4%	100%	<0.001
100%	0%	0%	0%	3%	0%	100%	100%	<0.001

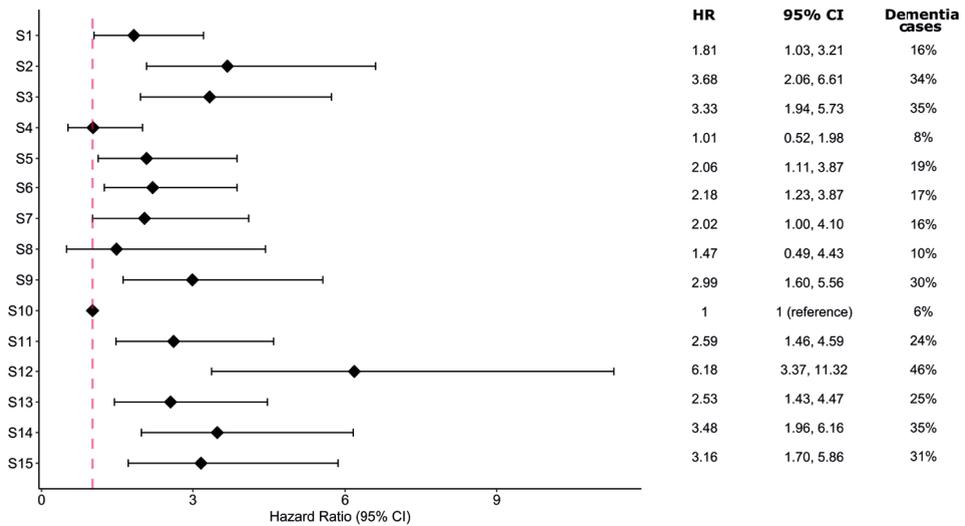


Figure 6.4. Hazard ratios per subgroup based on the Cox regression analysis. Controlled for age, sex, and cognitive status at baseline. Subgroup 10 was used as the reference subgroup in the model, as it was the subgroup with the lowest amount of brain abnormalities. S: subgroup; HR: hazard ratio; CI: confidence interval.

6.5 DISCUSSION

We showed that distinct brain MRI phenotypes can be identified in community-dwelling older adults. Some of the 15 distinct subgroups that were identified showed a different long-term dementia risk, with an increased risk, especially in individuals in the multi-burden brain pathology subgroups, and in the subgroup with relatively severe WMH and atrophy. Most subgroups showed a significantly increased risk for dementia compared with the reference subgroup that showed the least abnormalities on brain MRI.

The exact underlying structural correlates of the early predisposition to dementia remain largely unknown. Many validated and commonly used brain MRI markers are nonspecific to pathology and disease and commonly occur with ageing.¹⁹ Nevertheless, previous studies frequently focused on single or small groups of MRI markers and their relationship with neurovascular or neurodegenerative diseases.²⁰⁻²³ A combined analysis of brain MRI markers could improve our understanding of the pathophysiology of and early predisposition for dementia, as different brain diseases usually lead to patterns of different brain abnormalities.²⁴ In other fields of research, combined analysis to identify phenotypes has previously been performed. Examples are the identification of asthma phenotypes^{25,26} and subphenotypes of chronic

obstructive pulmonary disease.^{27,28} Moreover, it is frequently used in the field of genetics to identify genotypes, for example, to identify differences in DNA methylation and gene expression in breast cancer.²⁹

There are some previous studies on combined analysis of brain MRI markers or data-driven approaches in the field of dementia research. In a previous study, using an unsupervised deep learning approach on a large diverse data set of T1-weighted brain MRI scans showed that the difference between predicted brain age and chronological age is associated with the presence of different diseases (e.g., schizophrenia, and Alzheimer disease (AD)).³⁰ In another study, a semi-supervised deep-clustering method identified 4 neurodegenerative brain MRI patterns based on atrophy regions of interests on T1 scans in a data set including cognitively healthy individuals and patients with cognitive impairment and dementia.³¹ Another study identified 3 brain MRI patterns (neurodegeneration, white matter disease, and typical brain ageing) on T1-weighted scans using a machine learning-based method that can be used to identify individual brain health.³² These previous studies used (semi) supervised deep learning approaches, which is a different approach compared with our unsupervised machine learning approach. There are also some previous studies that have used a more similar approach compared with ours, albeit in different patient populations. For example, a previous study has applied hierarchical clustering to identify patterns of markers (including brain MRI markers, blood values, and CSF markers) related to the conversion from mild cognitive impairment to AD.³³ Here, 4 subgroups were identified with a different risk for conversion to AD.³³ The subgroup with the highest risk showed the most severe biomarker profile, for example, the highest WMH volumes, the lowest CSF amyloid beta, the highest CSF tau, and the lowest entorhinal cortical thickness.³³ Another previous study showed that midlife white matter textural properties were associated with future dementia risk.³⁴ A more heterogeneous normal appearing white matter intensity profile was associated with a higher WMH burden in the future, and a more heterogeneous intensity of normal appearing white matter was related to increased dementia risk. Another study found 2 distinct subgroups of mild cognitive impairment based on radiomics similarity networks.³⁵ Significant differences between the 2 mild cognitive impairment subgroups were found, among others, in the regional radiomics similarity networks of the hippocampus, temporal lobe, parahippocampal gyrus, and amygdala, as well as in the gray matter volume and cortical thickness. Furthermore, the 2 subgroups were significantly different from each other in clinical measures and the number of participants progressing to dementia within 3 years.³⁵ Our study is the first to apply an unsupervised machine learning approach in a large group of community-dwelling individuals to assess the association of brain MRI phenotypes and long-term

dementia risk. The large sample size of community-dwelling individuals in our study aids the generalizability of our results. Our study is further strengthened by the long follow-up time to the assessment of occurrence of dementia. Because of our study design, in the future, our method could help in assessing an individuals increased dementia risk at an early stage to determine patient prognosis in clinical practice and may aid in patient selection for future treatment studies.

A hierarchical clustering model allows combining of a spectrum of brain MRI markers and to find patterns in these data. We have previously applied the hierarchical clustering method in different data sets with different MRI markers to identify MRI phenotypes of the brain related to future stroke and mortality in patients with manifest arterial disease,⁷ as well as related to increased postoperative delirium risk in preoperative patients.³⁶ To the best of our knowledge, this study assessed brain MRI phenotypes in relation to long-term dementia risk in community-dwelling older adults. Our results identified 15 distinct subgroups of individuals with different distributions of brain MRI markers of neurodegenerative and neurovascular disease. The multi-burden group with the highest long-term risk for dementia (subgroup 12) does show markers of SVD, such as high WMH volumes and an irregular WMH shape, but includes only few individuals with brain infarcts. In addition, subgroup 12 showed the most severe cerebral atrophy, which may suggest that this subgroup has more underlying neurodegenerative pathology. Subgroup 2 has, similar to subgroup 12, high WMH volumes and an irregular WMH shape but also includes a high number of participants with subcortical, cerebellar, and cortical infarcts. Atrophy is less prominent in subgroup 2 compared with subgroup 12. Subgroup 15 may include mostly patients with large vessel disease, as WMH volumes and shape are only moderately abnormal, while all participants in this group have cortical infarcts. We showed that different brain MRI phenotypes, characterized by a distinct combination of brain MRI markers, predispose to occurrence of dementia and are related to different long-term dementia risks.

Strengths of our study include the use of multiple brain MRI markers in one framework, a large sample size, and a long follow-up period for dementia outcome. Furthermore, we mostly included markers that can be (semi) automatically detected on brain MRI scans (e.g., WMH volumes and brain atrophy) and the inclusion of novel brain MRI markers (such as WMH shape). An automated, unsupervised approach to identify groups was applied that allowed us to identify novel patterns of brain MRI markers. Limitations of this study might be the somewhat subjective cut-offs within the model, such as the dendrogram cut-off. However, to increase objectivity, we used the Dunn index and the heatmap to determine the cut-off for the number of

subgroups. Another limitation could be that the model is dependent on the selection of brain MRI markers that were included in the model. We therefore chose to include etiologically and prognostically relevant and validated brain MRI markers of which many can be quantified automatically. Moreover, the hierarchical clustering results could also be influenced by the choice of linkage method (e.g., Ward's, centroid), which is used to delineate the subgroups. We have chosen to use Ward's criteria as a linkage method because it generates subgroups with minimal within subgroup variance and to maximize the between-subgroup variance, which we deemed as most suitable for this data set and type of analysis. Another general limitation of the clustering method could be that in our sensitivity analyses, we showed that there is some dependency of the clustering results based on the number and selection of participants. For future research the Subtype and Stage Inference (SuStain) method, an unsupervised machine learning technique that identifies population subgroups with common patterns of disease progression could be an interesting approach.³⁷ SuStain could provide additional insights since it combines traditional clustering with disease progression modeling, but the effect of this approach on reproducibility is also of interest. Another limitation of this study could be that most neuroimaging research 1.5T MRI scanners are nowadays replaced with a 3T MRI system, which was not yet the case at the time of the data collection for our study. Nevertheless, we did successfully identify distinct brain MRI phenotypes based on our data set. In conclusion, distinct brain MRI phenotypes are related to varying long-term risks of developing dementia. Brain MRI phenotypes may assist in an improved understanding of the structural correlates of dementia predisposition. These findings may aid in the future to determine patient prognosis and for patient selection for future treatment studies.

6.6 ACKNOWLEDGEMENTS

The Age, Gene/Environment Susceptibility Reykjavik Study was supported by NIH contracts N01-AG-1-2100 and HHSN27120120022C, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). This work was supported by an Alzheimer Nederland grant (WE.03-2019-08) to Jeroen de Bresser.

6.7 REFERENCES

1. Duering M, Biessels GJ, Brodtmann A, et al. Neuroimaging standards for research into small vessel disease—advances since 2013. *Lancet Neurol.* 2023;22(7):602-618. doi: 10.1016/S1474-4422(23)00131-X
2. Alber J, Alladi S, Bae H, et al. White matter hyperintensities in vascular contributions to cognitive impairment and dementia (VCID): knowledge gaps and opportunities. *Alzheimers Dement (NY).* 2019;5:107-117. doi:10.1016/j.trci.2019.02.001
3. Benveniste H, Nedergaard M. Cerebral small vessel disease: a glymphopathy? *Curr Opin Neurobiol.* 2022;72:15-21. doi:10.1016/J.CONB.2021.07.006
4. Bos D, Wolters FJ, Darweesh SKL, et al. Cerebral small vessel disease and the risk of dementia: a systematic review and meta-analysis of population-based evidence. *Alzheimers Dement.* 2018;14(11):1482-1492. doi:10.1016/j.jalz.2018.04.007
5. De Guio F, Duering M, Fazekas F, et al. Brain atrophy in cerebral small vessel diseases: extent, consequences, technical limitations and perspectives: the HARNESS initiative. *J Cereb Blood Flow Metab.* 2020;40(2):231-245. doi:10.1177/0271678X19888967
6. Ding J, Sigurdsson S, Jonsson PV, et al. Space and location of cerebral microbleeds, cognitive decline, and dementia in the community. *Neurology.* 2017;88(22):2089-2097. doi:10.1212/WNL.0000000000003983
7. Jaarsma-Coes MG, Ghaznawi R, Hendrikse J, et al; Second Manifestations of ARterial disease SMART Study group. MRI phenotypes of the brain are related to future stroke and mortality in patients with manifest arterial disease: the SMART-MR study. *J Cereb Blood Flow Metab.* 2020;40(2):354-364. doi:10.1177/0271678X18818918
8. Harris TB, Launer LJ, Eiriksdottir G, et al. Age, gene/environment susceptibility Reykjavik study: multidisciplinary applied phenomics. *Am J Epidemiol.* 2007;165(9):1076-1087. doi:10.1093/aje/kwk115
9. Saczynski JS, Sigurdsson S, Jonsdottir MK, et al. Cerebral infarcts and cognitive performance: importance of location and number of infarcts. *Stroke.* 2009;40(3): 677-682. doi:10.1161/STROKEAHA.108.530212
10. Morris JN, Hawes C, Fries BE, et al. Designing the national resident assessment instrument for nursing homes. *Gerontologist.* 1990;30(3):293-307. doi:10.1093/GERONT/30.3.293
11. Sveinbjornsdottir S, Sigurdsson S, Aspelund T, et al. Cerebral microbleeds in the population based AGES-Reykjavik study: prevalence and location. *J Neurol Neurosurg Psychiatry.* 2008;79(9):1002-1006. doi:10.1136/JNNP.2007.121913
12. Zijdenbos AP, Forghani R, Evans AC. Automatic “pipeline” analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. *IEEE Trans Med Imaging.* 2002;21(10):1280-1291. doi:10.1109/TMI.2002.806283
13. Sigurdsson S, Aspelund T, Forsberg L, et al. Brain tissue volumes in the general population of the elderly: the AGES-Reykjavik Study. *Neuroimage.* 2012;59(4):3862-3870. doi:10.1016/j.neuroimage.2011.11.024
14. Keller JA, Sigurdsson S, Klaassen K, et al. White matter hyperintensity shape is associated with long-term dementia risk. *Alzheimers Dement.* 2023;19(12):5632-5641. doi:10.1002/ALZ.13345
15. Ghaznawi R, Geerlings MI, Jaarsma-Coes MG, et al. The association between lacunes and white matter hyperintensity features on MRI: the SMART-MR study. *J Cereb Blood Flow Metab.* 2019;39(12):2486-2496. doi:10.1177/0271678X18800463
16. Kassambara A. FactoExtra: extract and visualize the results of multivariate data analyses for R, version 1.0.7. 2020.
17. Maechler M, Rousseeuw P, Struyf A, et al. Package “cluster”: R package version 2.1.2.2021.

18. Galili T. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*. 2015;31(22):3718-3720. doi:10.1093/bioinformatics/btv428
19. Pantoni L. Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges. *Lancet Neurol*. 2010;9(7):689-701. doi:10.1016/S1474-4422(10)70104-6
20. Apostolova LG, Mosconi L, Thompson PM, et al. Subregional hippocampal atrophy predicts Alzheimer's dementia in the cognitively normal. *Neurobiol Aging*. 2010;31(7):1077-1088. doi:10.1016/J.NEUROBIOLAGING.2008.08.008
21. Debette S, Beiser A, Decarli C, et al. Association of MRI markers of vascular brain injury with incident stroke, mild cognitive impairment, dementia, and mortality: the Framingham Offspring Study. *Stroke*. 2010;41(4):600-606. doi:10.1161/STROKEAHA.109.570044
22. Godin O, Tzourio C, Rouaud O, et al. Joint effect of white matter lesions and hippocampal volumes on severity of cognitive decline: the 3C-Dijon MRI study. *J Alzheimers Dis*. 2010;20(2):453-463. doi:10.3233/JAD-2010-1389
23. van Rooden S, Goos JDC, van Opstal AM, et al. Increased number of microinfarcts in Alzheimer disease at 7-T MR imaging. *Radiology*. 2014;270(1):205-211. doi:10.1148/radiol.13130743
24. Habes M, Grothe MJ, Tunc B, McMillan C, Wolk DA, Davatzikos C. Disentangling heterogeneity in Alzheimer's disease and related dementias using data-driven methods. *Biol Psychiatry*. 2020;88(1):70-82. doi:10.1016/J.BIOPSYCH.2020.01.016
25. Haldar P, Pavord ID, Shaw DE, et al. Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med*. 2008;178(3):218-224. doi:10.1164/RCCM.200711-1754OC
26. Moore WC, Meyers DA, Wenzel SE, et al; National Heart, Lung, and Blood Institute's Severe Asthma Research Program. Identification of asthma phenotypes using cluster analysis in the Severe Asthma Research Program. *Am J Respir Crit Care Med*. 2010;181(4):315-323. doi:10.1164/RCCM.200906-0896OC
27. Burgel PR, Paillasser JL, Caillaud D, et al; Initiatives BPCO Scientific Committee. Clinical COPD phenotypes: a novel approach using principal component and cluster analyses. *Eur Respir J*. 2010;36(3):531-539. doi:10.1183/09031936.00175109
28. Fens N, Van Rossum AGJ, Zanen P, et al. Subphenotypes of mild-to-moderate COPD by factor and cluster analysis of pulmonary function, CT imaging and breathomics in a population-based survey. *COPD*. 2013;10(3):277-285. doi:10.3109/15412555.2012.744388
29. Lin IH, Chen DT, Chang YF, et al. Hierarchical clustering of breast cancer methylomes revealed differentially methylated and expressed breast cancer genes. *PLoS One*. 2015;10(2):e0118453. doi:10.1371/JOURNAL.PONE.0118453
30. Bashyam VM, Erus G, Doshi J, et al. MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide. *Brain*. 2020;143(7):2312-2324. doi:10.1093/BRAIN/AWAA160
31. Yang Z, Nasrallah IM, Shou H, et al; iSTAGING Consortium; Baltimore Longitudinal Study of Aging BLSA; Alzheimer's Disease Neuroimaging Initiative ADNI. A deep learning framework identifies dimensional representations of Alzheimer's Disease from brain structure. *Nat Commun*. 2021;12:7065. doi: 10.1038/s41467-021-26703-z
32. Habes M, Pomponio R, Shou H, et al; iSTAGING consortium, the Preclinical AD consortium, the ADNI, and the CARDIA studies. The Brain Chart of Aging: machine-learning analytics reveals links between brain aging, white matter disease, amyloid burden, and cognition in the iSTAGING consortium of 10,216 harmonizedMRscans. *Alzheimers Dement*. 2021;17(1):89-102. doi:10.1002/ALZ.12178
33. Nettiksimmons J, DeCarli C, Landau S, Beckett L; Alzheimer's Disease Neuroimaging Initiative. Biological heterogeneity in ADNI amnesic mild cognitive impairment. *Alzheimers Dement*. 2014;10(5):511-521.e1. doi:10.1016/J.JALZ.2013.09.003

34. Dounavi ME, Low A, Muniz-Terrera G, et al. Fluid-attenuated inversion recovery magnetic resonance imaging textural features as sensitive markers of white matter damage in midlife adults. *Brain Commun.* 2022;4(3):fcac116. doi:10.1093/BRAINCOMMS/FCAC116
35. Zhao K, Zheng Q, Dyrba M, et al; Alzheimer's Disease Neuroimaging Initiative. Regional radiomics similarity networks reveal distinct subtypes and abnormality patterns in mild cognitive impairment. *Adv Sci.* 2022;9(12):e2104538. doi:10.1002/ADVS.202104538
36. Kant IMJ, Sooter AJC, Jaarsma-Coes M, et al; BioCog consortium. Preoperative MRI brain phenotypes are related to postoperative delirium in older individuals. *Neurobiol Aging.* 2021;101:247-255. doi:10.1016/J.NEUROBIOLAGING.2021.01.033
37. Young AL, Marinescu RV, Oxtoby NP, et al; Genetic FTD Initiative GENFI; Alzheimer's Disease Neuroimaging Initiative ADNI. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. *Nat Commun.* 2018;9(1):4273. doi:10.1038/S41467-018-05892-0

6.8 SUPPLEMENTARY MATERIAL

Table S.6.8.1. Definition of the WMH shape markers.

Shape marker	WMH type	Description	Formula
Solidity (S)	Periventricular/Confluent WMH	Solidity and Convexity show how concave or convex a shape is. A maximally convex shape has a convexity and solidity value of 1. Values decrease with a more concave, complex shape. A convex hull is the smallest convex envelop of a shape.	$S = \frac{\text{Volume}}{\text{Convex Hull Volume}}$
Convexity (C)	Periventricular/Confluent WMH		$C = \frac{\text{Convex Hull Area}}{\text{Area}}$
Concavity Index (CI)	Periventricular/Confluent WMH	Concavity index—a measure of roughness—describes how dense, irregular or elongated and curved a lesion is. Higher CI values indicate a more complex WMH shape.	$CI = \sqrt{(2 - C)^2 + (1 - S)^2}$
Fractal Dimension (FD)	Periventricular/Confluent and deep WMH	The Minkowski-Bouligand dimension (box counting dimension) is a measure of textural roughness. Higher FD values suggest a more complex WMH shape.	$FD = \lim_{r \rightarrow 1} \frac{\log(n_r)}{\log\left(\frac{1}{r}\right)}$
Eccentricity (E)	Deep WMH	Eccentricity describes the deviation from a circle. The eccentricity of a circle is 0 and the eccentricity of a line is 1.	$E = \frac{\text{Major Axis}}{\text{Minor Axis}}$ <p>n = number of boxes r = box size</p>

Major axis: largest diameter in 3D space.
Minor axis: smallest diameter orthogonal to the major axis.

Lower convexity and solidity and a higher concavity index and fractal dimension indicate a more irregularly shaped WMH. A higher eccentricity corresponds to a more elongated lesion, while a lower eccentricity corresponds to a rounder lesion. Averages were calculated per participant for each shape marker. A more irregular shape of a WMH type was defined by taking the separate markers and their directionality into account to assess the main shape. This was only done for interpretation and description of the findings.

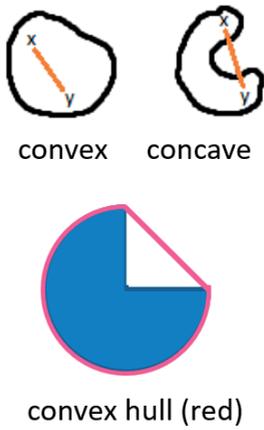


Figure S.6.8.1. Simplified illustrations of the concepts of convexity and convex hull. A shape is convex if you can connect any two points (e.g. x and y) within the shape, while the connecting line (orange) is also always within the shape. If this is not possible the shape is concave. A convex hull (red) is the smallest convex set that contains the shape (blue).

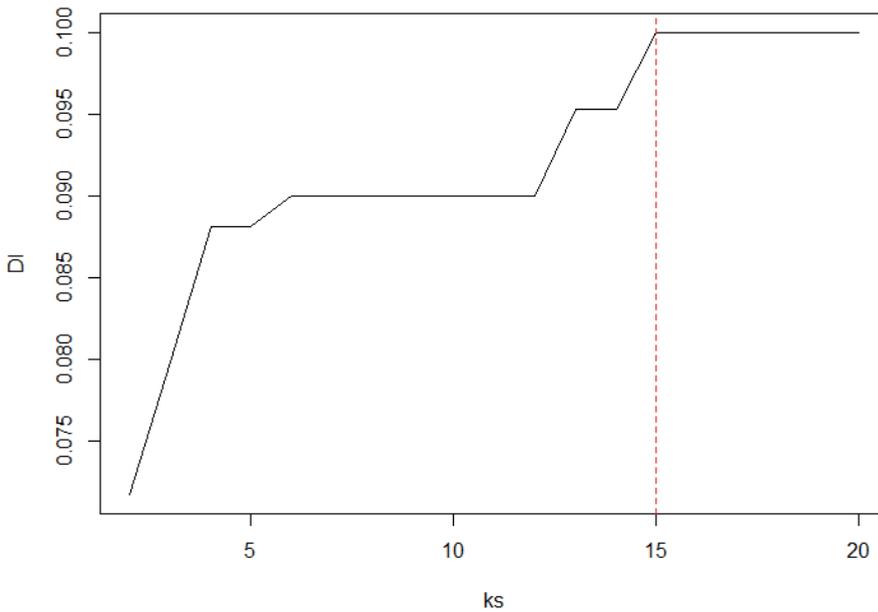


Figure S.6.8.2. The Dunn index is shown on the y-axis, and on the x-axis the number of clusters/subgroups is shown. The Dunn index should be maximized for an optimal number of groups. The red line indicates 15 groups, which is the number of groups used in the analysis. ks, number of clusters; DI, Dunn index.

Table S.6.8.2. Between-group differences of WMH volumes per lobe.

	S1 N=368	S2 N=161	S3 N=425	S4 N=239	S5 N=171	S6 N=333	S7 N=98
PV/C WMH volume	2.84 ±	12.89 ±	10.33 ±	1.86 ±	3.16 ±	1.74 ±	1.23 ±
frontal lobe	1.66	12.26	9.30	1.34	3.24	1.07	0.75
PV/C WMH volume	1.78 ±	8.06 ±	6.75 ±	0.78 ±	1.76 ±	0.98 ±	0.44 ±
parietal lobe	1.63	6.48	5.42	0.93	2.06	0.89	0.45
PV/C WMH volume	2.16 ±	5.10 ±	4.44 ±	1.17 ±	1.89 ±	1.54 ±	0.98 ±
temporal lobe	1.15	2.74	2.23	0.69	1.34	0.87	0.53
PV/C WMH volume	1.68 ±	3.35 ±	3.25 ±	0.55 ±	1.42 ±	1.71 ±	0.85 ±
occipital lobe	1.01	2.41	1.73	0.59	1.26	1.03	0.76
Other PV/C WMH	2.23 ±	4.36 ±	3.74 ±	1.69 ±	2.04 ±	1.60 ±	1.33 ±
volume	0.75	2.02	1.41	0.66	0.98	0.65	0.55
Deep WMH volume	0.79 ±	1.44 ±	0.98 ±	0.46 ±	0.79 ±	0.24 ±	0.27 ±
frontal lobe	0.72	1.20	0.71	0.50	1.09	0.28	0.36
Deep WMH volume	0.21 ±	0.33 ±	0.22 ±	0.11 ±	0.21 ±	0.07 ±	0.07 ±
parietal lobe	0.25	0.34	0.19	0.13	0.34	0.09	0.13
Deep WMH volume	0.03 ±	0.07 ±	0.04 ±	0.06 ±	0.05 ±	0.02 ±	0.03 ±
temporal lobe	0.05	0.09	0.05	0.09	0.08	0.03	0.05
Deep WMH occipital	0.10 ±	0.11 ±	0.09 ±	0.74 ±	0.37 ±	0.09 ±	0.52 ±
lobe	0.17	0.20	0.18	0.47	0.52	0.14	0.60
Other deep WMH	0.05 ±	0.13 ±	0.09 ±	0.08 ±	0.08 ±	0.03 ±	0.04 ±
volume	0.06	0.12	0.10	0.09	0.10	0.04	0.06

Data are shown as means ± SD. Brain MRI markers were compared between subgroups with one-way ANOVAs. S, subgroup; WMH, white matter hyperintensity; PV/C periventricular/confluent WMH; SD, standard deviation. Other PV/C WMH volume or other deep WMH volume are defined as lesions or parts of lesions that outside of the brain lobe masks, towards the brain stem, and internal capsule.

S8 N=42	S9 N=105	S10 N=241	S11 N=240	S12 N=87	S13 N=245	S14 N=190	S15 N=111	P value
1.35 ± 0.81	6.99 ± 6.30	0.71 ± 0.46	8.35 ± 7.36	18.00 ± 11.01	7.28 ± 5.08	15.71 ± 12.19	6.91 ± 7.89	<0.000
0.66 ± 0.78	4.55 ± 3.84	0.16 ± 0.23	± 5.15	10.91 ± 6.03	4.80 ± 3.27	9.50 ± 6.50	4.79 ± 5.31	<0.000
1.11 ± 0.68	3.61 ± 2.04	0.46 ± 0.35	3.93 ± 2.26	7.15 ± 3.13	4.03 ± 1.85	5.66 ± 2.68	3.66 ± 2.33	<0.000
0.75 ± 0.65	2.62 ± 1.58	0.21 ± 0.32	2.69 ± 1.77	5.54 ± 3.14	1.83 ± 1.07	3.80 ± 2.31	2.87 ± 2.02	<0.000
1.31 ± 0.58	3.28 ± 1.38	0.84 ± 0.43	3.61 ± 1.58	5.39 ± 1.69	3.49 ± 1.04	4.53 ± 1.65	3.05 ± 1.58	<0.000
0.27 ± 0.36	1.27 ± 1.05	0.10 ± 0.14	1.36 ± 0.96	2.11 ± 1.11	2.14 ± 1.43	1.30 ± 1.05	1.03 ± 1.14	<0.000
0.07 ± 0.08	0.32 ± 0.32	0.03 ± 0.04	0.31 ± 0.26	0.60 ± 0.50	0.56 ± 0.47	0.34 ± 0.33	0.26 ± 0.24	<0.000
0.04 ± 0.05	0.06 ± 0.09	0.03 ± 0.04	0.07 ± 0.09	0.29 ± 0.29	0.11 ± 0.14	0.07 ± 0.09	0.04 ± 0.06	<0.000
0.47 ± 0.44	0.09 ± 0.15	0.64 ± 0.48	0.13 ± 0.23	0.11 ± 0.11	0.18 ± 0.34	0.09 ± 0.12	0.17 ± 0.27	<0.000
0.05 ± 0.07	0.11 ± 0.11	0.04 ± 0.05	0.14 ± 0.14	0.22 ± 0.14	0.14 ± 0.14	0.12 ± 0.11	0.07 ± 0.09	<0.000

S.6.8.1 Supplementary results

Our model resulted in 15 distinct subgroups with unique combinations of brain MRI markers. Below is a more detailed description of the brain MRI markers and the most important characteristics per subgroup. Subjective and relative cut-off points were used for the below used description of the subgroups. For continuous variables the subgroups were split into three thirds, with the highest, moderate and lowest values. WMH shape was considered as regular/moderate/irregular looking at all shape markers together and assessing the main shape. A low solidity, low convexity, high concavity index, and high fractal dimension indicate a more irregular shape. A high eccentricity indicates a rounder shape, while a low eccentricity indicates a more elongated shape. Total WMH volumes were used to describe relatively low/moderate/high WMH volumes. WMH below ~20 ml were considered relatively low volumes, between ~20 ml and ~30 ml relatively moderate volumes and above ~30 ml as relatively high volumes. Total brain volumes above ~74% of intracranial volume were considered as relatively low atrophy, between ~72 and ~74% of intracranial volume as relatively moderate atrophy and below ~72% of intracranial volume as relatively severe atrophy.

Subgroup 1 has relatively low WMH volumes, no infarcts, microbleeds or enlarged PVS, a moderately irregular WMH shape and relatively low atrophy.

Subgroup 2 has relatively high WMH volumes, a high prevalence of subcortical infarcts (100%), some cortical, and cerebellar infarcts, an irregular WMH shape, some enlarged PVS (in and around the basal ganglia) and microbleeds, and relatively severe cerebral atrophy.

Subgroup 3 has relatively high WMH volumes, some enlarged PVS (in and around the basal ganglia), an irregular WMH shape, and relatively severe cerebral atrophy.

Subgroup 4 has relatively low WMH volumes, no infarcts, almost no enlarged PVS, a relatively regular WMH shape, some microbleeds, and relatively low cerebral atrophy.

Subgroup 5 has relatively low WMH volumes, cerebellar infarcts (98%), some enlarged PVS (in and around the basal ganglia, and white matter) and microbleeds, a moderately irregular WMH shape, and moderate cerebral atrophy.

Subgroup 6 has relatively low WMH volumes, a moderately irregular WMH shape, no infarcts, no enlarged PVS or microbleeds, and moderate cerebral atrophy.

Subgroup 7 has relatively low WMH volumes, no infarcts, a relatively regular WMH shape, enlarged PVS (in the white matter 100%), no microbleeds, and moderate cerebral atrophy.

Subgroup 8 has relatively low WMH volumes, a relatively regular WMH shape, enlarged PVS (mostly in and around the basal ganglia, and some in the white matter), a high prevalence of subcortical infarcts (100%) and a relatively low amount of cerebral atrophy.

Subgroup 9 has moderately high WMH volumes, a relatively irregular WMH shape, the most microbleeds (66%), no infarcts, some enlarged PVS (in the white matter), and moderate cerebral atrophy.

Subgroup 10 has the lowest WMH volumes, no infarcts or enlarged PVS, the most regular WMH shape, and the lowest amount of cerebral atrophy (the highest brain volumes). Furthermore, this subgroup has the lowest number of participants with hypertension (64%), type 2 diabetes (4%), and includes the most female participants (81%).

Subgroup 11 has moderately high WMH volumes, a relatively irregular WMH shape, no infarcts, some microbleeds, enlarged PVS (mostly in the white matter, 99%; basal ganglia: 8%), and moderate cerebral atrophy.

Subgroup 12 has the highest WMH volumes, some subcortical infarcts, some PVS (in and around the basal ganglia, and whole brain), the most irregular WMH shape, and the most severe cerebral atrophy. Moreover, this subgroup has the highest deep WMH volumes in the temporal lobe (0.29 ± 0.29 ml) and the highest percentage of participants with hypertension (90%).

Subgroup 13 has moderately high WMH volumes, no infarcts or enlarged PVS, a relatively irregular WMH shape, and a relatively low amount of cerebral atrophy.

Subgroup 14 has relatively high WMH volumes and a relatively irregular WMH shape, some enlarged PVS (in and around the basal ganglia, and in the white matter), some subcortical infarcts, a large amount of cerebellar infarcts (99%), some cortical infarcts, and a relatively high amount of cerebral atrophy.

Subgroup 15 has moderately high WMH volumes, a moderately irregular WMH shape, some cerebellar infarcts (29%), a large amount of cortical infarcts (100%), some subcortical infarcts, a relatively high amount of cerebral atrophy, some microbleeds, and some enlarged PVS (in and around the basal ganglia, and in the white matter).