# Permutation-based inference preserves anatomical specificity in lesion network mapping

Marvin Petersen[1], Kaustubh R. Patil[2,3], Simon B. Eickhoff[2,3], Geert Jan Biessels[1] on behalf of the Meta VCI Map Consortium

## Abstract

Because many neurobehavioural functions rely on distributed brain networks, anatomically diverse brain lesions can cause the same neurobehavioural deficit. Lesion network mapping (LNM) builds on this principle to better understand the functional anatomy of the brain, by mapping focal lesions onto a normative functional connectome. Yet, recent work raised concerns about anatomical specificity of LNM, showing that commonly used LNM procedures converge on nonspecific connectome properties. Here, we show that anatomically specific LNM is possible with the right statistical approach – using symptom-label permutation as a null model. We demonstrate this in a multicenter dataset of 2,950 stroke patients across 12 cohorts, comparing patients with and without impairment in 6 cognitive domains. First, we showed that permutation-based LNM yielded distinct and biologically plausible network maps with modest cross-cognitive domain similarity. Second, we replicated the previously raised concern of nonspecific connectome-driven convergence when using parametric statistics. Third, we assessed specificity of our approach through simulation analyses across 10,000 null studies which confirmed that the permutation framework maintained valid type I error control. These findings demonstrate that permutation-based null models preserve anatomical specificity in LNM, enabling the identification of brain networks that are genuinely linked to distinct neurobehavioural functions. This approach may thus allow researchers to more reliably map the network basis of neurobehavioural deficits from focal brain lesions.

**Author affiliations:**

1 University Medical Center Utrecht Brain Center, Utrecht, The Netherlands

2 Institute for Systems Neuroscience, Medical Faculty, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany

3 Institute of Neuroscience and Medicine, Brain and Behaviour (INM-7), Research Center Ju□lich, Germany

Correspondence to: m.petersen-28@umcutrecht.nl; g.j.biessels@umcutrecht.nl

**Running title:** Permutation-based lesion network mapping

**Keywords:** lesion network mapping; permutation; connectome; post-stroke cognitive impairment; vascular cognitive impairment; neuroimaging

# Introduction

Why anatomically distinct lesions can cause the same neurobehavioural symptom has been a longstanding enigma in neuroscience. One explanation is that such distinct lesions may share membership in a common brain network: a lesion anywhere in a network fundamental to a given function will impair that function, regardless of the exact lesion site. Lesion network mapping (LNM) operationalizes this principle, providing an analytical framework for understanding why distinct lesions can produce similar symptoms and providing insight into the functional anatomy of the brain from a network perspective.[1] By projecting lesion locations on a normative functional connectome, LNM seeks to uncover the distributed neural circuits underlying clinical deficits.[2] This promise has contributed to its rapid adoption across neurology and psychiatry.

However, recent work has raised concerns about the technique.[3] Across many studies, lesion network maps appear strikingly similar even for clinically distinct symptoms. This implausible convergence stems from a methodological bias: common LNM procedures tend to highlight generic features of brain organization – such as hub connectivity (degree) – by repeatedly sampling the fixed normative connectome.[3] As a result, LNM of clinically unrelated conditions such as addiction, epilepsy and psychosis can yield nearly identical network maps. This convergence appears to affect LNM regardless of study design, including approaches that only average connectivity maps across patients as well as variants that statistically link patient-level connectivity and clinical information through group contrasts or correlations with continuous symptom scores.[3] Together, these findings have cast doubt on whether LNM can capture genuine symptom-linked biology, calling for a first-principles revision.

This critique prompted us to revisit our own earlier work, published in this Journal.[4] We found that unlike many previous studies,[3] our LNM implementation already accounts for this bias through its use of symptom-label permutation. Symptom-label permutation preserves the structured dependence induced by the connectome while breaking the relationship between symptoms and imaging features. Because shuffling symptom labels reproduces the convergence phenomenon in every permutation, the nonspecific effect becomes part of the null distribution. Under this strict null, only signals exceeding nonspecific connectome properties remain significant. Crucially, this approach permutes symptom labels – not lesion locations – so the lesion-derived connectivity maps remain fixed while only the symptom–imaging coupling is disrupted. We initially implemented this approach because inference based on symptom-label permutation is well established in neuroimaging and it has also been adopted in other LNM studies.[5–7] However, if it can overcome the observed non-specific connectome-driven convergence has not been systematically examined.

Here, we aim to establish that, given the use of appropriate techniques, such as symptom-label permutation, anatomically specific LNM is achievable – implying that lesion network maps can capture brain networks genuinely linked to distinct neurobehavioural functions, rather than reflecting nonspecific connectome properties. To demonstrate this, we systematically evaluate symptom-label permutation, showing that it does produce distinct network maps for different cognitive domains, that are anatomically plausible. We also

compare it to parametric voxelwise testing, which has been used in many previous LNM studies.[3] Parametric testing implicitly assumes independence among lesion-derived connectivity maps. Yet, these maps substantially covary because they are all derived from the same fixed connectome, making conventional statistical inference vulnerable to false positives and consequently to inflated anatomical similarity. We identify this as a key contributor to the observed convergence.

Using a multicenter dataset of 2,950 stroke patients, we show that permutation-based inference mitigates connectome-driven convergence and yields anatomically specific network maps.

# Materials and methods

We analyzed data from 2,950 stroke patients across 12 post-stroke cognitive impairment cohorts of the Meta VCI Map Consortium, all with acute ischemic stroke lesion segmentations and harmonized cognitive assessment.[8] Cognitive status was classified as impaired or unimpaired for six domains – attention/executive function, processing speed, language, verbal memory, visuospatial function, and visuospatial memory – based on performance below the fifth percentile of local normative data. Cognitive impairment labels provided the basis for all statistical comparisons.

Acute lesions were segmented within each cohort and normalized to MNI152NLin6Asym space. These normalized lesion masks were used to compute lesion-derived functional connectivity maps, which estimate the functional networks perturbed by each patient's lesion. The connectivity maps were computed with Lead-DBS using the GSP1000 normative functional connectome.[9,10]

Symptom-linked LNM was performed using a voxelwise general linear model that compared lesion-derived connectivity maps between cognitively impaired and unimpaired groups. Two inference strategies were compared within this model. The first derived p-values parametrically from the model's t-statistics. The second used symptom-label permutation,[11] generating null distributions by randomly reassigning impairment labels while keeping connectivity maps and covariates fixed. Here, p-values were defined as the proportion of permuted test statistics exceeding the observed statistic.

To quantify pairwise similarity across cognitive domains, we computed spatial correlations of T-statistic maps and Dice overlap of thresholded significance maps. Pairwise similarity with the degree map of the GSP1000 connectome was also quantified. To determine whether permutation-based LNM captures information beyond focal lesion location, we further compared the resulting networks with voxel-based lesion-symptom mapping performed on the same dataset.

Statistical analyses were implemented in Permutation Analysis of Linear Models (PALM).[11] Lesion volume was included as a covariate, voxelwise statistics were corrected for multiple comparisons based on false-discovery-rate, and permutation-based testing employed 1,000 permutations with threshold-free cluster enhancement.

To assess type I error control of the permutation framework under the null hypothesis, we conducted simulation analyses across 10,000 null studies using both synthetic lesions and real lesion masks from the Meta VCI Map cohort. Under a valid null, the proportion of significant ROIs ($|Z| > 1.96$) should approximate the nominal false positive rate (≈5%). We quantified the empirical false positive rate and characterized its spatial distribution and associations with nonspecific connectome properties (degree and PC1). Details are presented in *supplementary figure S3*.

Each contributing cohort obtained local ethics approval and written informed consent, as previously described.[8,12]

# Results

When statistical significance was assessed against label-permutation nulls the resulting lesion network maps were anatomically specific (***Figure 1***). The analyses show that the permutation-based maps showed low spatial overlap with the connectome's degree map (mean Dice = 0.18 ± 0.08), implying that the identified networks were not driven by nonspecific connectome structure. They also exhibited only moderate similarity across different cognitive domains (mean Dice = 0.45 ± 0.14), indicating that the maps did not converge onto a shared, domain-general pattern.

This anatomical specificity was most readily apparent for language. The permutation-based language network preferentially highlighted left-lateralized perisylvian regions, including left inferior frontal and temporal cortex, aligning with a priori expectations for language circuitry. Visuospatial memory showed no significant group differences when evaluated against the label permutation null, suggesting that the observed connectivity patterns associated with visuospatial memory impairment did not exceed what would be expected from nonspecific connectome properties alone.

In contrast, in the same dataset the parametric LNM approach generated maps that corresponded to a higher extent with the normative degree map than (mean r = 0.58 ± 0.05; mean Dice = 0.30 ± 0.01) and showed biologically implausible high similarity across the different cognitive domains (mean r = 0.94 ± 0.03; mean Dice = 0.88 ± 0.05) (***Figure 2***). For example, the parametric language map showed limited specificity, without the expected left lateralization, with substantial non-specific alignment to the normative degree map. Overlap between parametric and permutation-based significance maps was moderate (mean Dice = 0.53 ± 0.15) (***Supplementary figure 1***).

Permutation-based LNM maps only partially overlapped with voxel-based lesion-symptom mapping (VLSM) maps for the same cognitive outcomes (mean Dice = 0.39 ± 0.07) (***Supplementary figure 2***), indicating that permutation-based LNM captures network-level information beyond what focal lesion location alone can explain.

Simulation analyses confirmed valid type I error control under the permutation null (***Supplementary figure 3***): across 10,000 null studies – combining random lesion masks with randomly assigned symptom labels – the observed mean false positive rate was 5.0%, closely matching the nominal rate. Permutation correction markedly reduced map alignment with connectome degree and PC1. Uncorrected maps, by contrast, showed stronger alignment with

degree and connectome PC1. These results were consistent across both real and synthetic lesion conditions.

Together, these results indicate that a sizeable fraction of the aberrant cross-domain convergence observed under parametric inference does not occur once inference is grounded in a null model that preserves connectome-induced dependence. As such, permutation-based inference can yield higher apparent domain specificity.

# Discussion

We demonstrate that symptom-label permutation yields anatomically specific and biologically plausible LNM results in a large multicenter stroke dataset. By contrast, if we apply parametric LNM on the same dataset, we replicate the connectome-driven convergence identified as an important source of bias as identified in previous work.[3] Therefore, by providing a viable methodological alternative, we demonstrate how LNM inference can be strengthened. Label permutation preserves the connectome-induced covariance in both the observed maps and the null, while selectively breaking symptom-imaging coupling, so that surviving effects are expected to be smaller but more anatomically specific. Importantly, this logic applies whenever a clinical variable anchors the analysis – whether a group contrast as used here, continuous symptom scores, or treatment outcomes – but does not extend to LNM designs that aggregate connectivity maps without such a reference.

The persistence of nontrivial maps for several domains under the permutation-based null supports the presence of genuine symptom-linked network effects beyond nonspecific connectome properties. These results align with our work showing that LNM significantly improves individual-level symptom prediction compared to pure lesion location – something that would not be possible if patient-level lesion connectivity maps carried only nonspecific information.[4] By contrast, the absence of significant effects for visuospatial memory underscores that not all outcomes remain detectable once stringent control is imposed. This likely reflects a trade-off in which anatomical specificity is gained at the cost of reduced sensitivity for weaker effects. Simulation analyses corroborated this interpretation: under the label-permutation null – where no true lesion–symptom relationship exists – the permutation framework controlled type I error at the nominal 5% level across 10,000 runs, while uncorrected maps showed stronger spurious alignment with connectome hubs. This confirms that the specificity gains observed in the empirical analyses reflect removal of nonspecific connectome-driven signal – although for domains with weaker or more diffuse effects, such as visuospatial memory, genuine signals may fall below the stricter detection threshold.

On this basis, we recommend incorporating symptom-label permutation or comparable non-parametric approaches that account for connectome-induced dependence, in LNM inference. Such methods provide a principled safeguard against connectome-driven convergence and sharpen what can be interpreted as a genuine disease-specific biological brain network. Framed this way, the identified limitation of LNM is not merely cautionary but actionable and a first-principles revision of the method may be closer at hand than expected.

## Data availability

The domain-level parametric and permutation-based lesion network maps are provided on Open Science Framework (https://osf.io/puq36). The individual-level data that support the findings of this study are available from the project leads on reasonable request (https://metavcimap.org/group/become-a-member/). Restrictions related to privacy and personal data sharing regulations and informed consent may apply. The analysis code is provided on GitHub (https://github.com/umcu-VCI-group/2026_petersen-revisiting_lesion_network_mapping).


## Funding

This project is part of the Timely, Accurate, and Personalized Diagnosis of Dementia (TAP-dementia) program, which receives funding from ZonMw (#10510032120003) in the context of Onderzoeksprogramma Dementie, which is part of the Dutch National Dementia Strategy. Further funding comes from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 579029650 (M.P.).


## Competing interests

The authors report no competing interests.

## Author contributions

Each author has made a significant contribution to the manuscript and all authors read and approved its final version. M.P.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Visualization, Writing—original draft, Writing – review & editing; K.P.: Investigation, Writing—review & editing; S.B.E.: Investigation, Writing – review & editing; G.J.B.: Conceptualization, Funding, Investigation, Resources, Supervision, Writing – review & editing. Contributors from the Meta VCI Map Consortium are listed in Appendix 1.

## Supplementary material

Supplementary material is available at *Brain* online.

## Appendix 1

The authors acknowledge the contributions of all researchers who provided resources to the Meta VCI Map post-stroke cognitive impairment initiative: Nick Weaver, PhD; J Matthijs Biesbroek, PhD; Hugo J Kuijf, PhD; Hugo P Aben, MD; Jill Abrigo, MD; Prof Hee-Joon Bae, PhD; Mélanie Barbay, MD; Jonathan G Best, MD; Prof Régis Bordet, PhD; Francesca M Chappell, PhD; Christopher P L H Chen, MD; Prof Martin Dichgans, MD; Thibaut

Dondaine, PhD; Jule Filler; Ruben S van der Giessen, PhD; Prof Olivier Godefroy, PhD; Bibek Gyanwali, MD; Olivia K L Hamilton, MSc; Saima Hilal, PhD; Irene M C Huenges Wajer, PhD; Prof Yeonwook Kang, PhD; Prof L Jaap Kappelle, PhD; Beom Joon Kim, PhD; Sebastian Köhler, PhD; Paul L M de Kort, PhD; Prof Peter J Koudstaal, PhD; Gregory Kuchcinski, MD; Bonnie Y K Lam, PhD; Prof Byung-Chul Lee, PhD; Keon-Joo Lee, MD; Jae-Sung Lim, MD; Renaud Lopes, PhD; Stephen D J Makin, PhD; Anne-Marie Mendyk, RN; Prof Vincent C T Mok, MD; Mi Sun Oh, MD; Prof Robert J van Oostenbrugge, PhD; Martine Roussel, PhD; Lin Shi, PhD; Julie Staals, PhD; Maria del C Valdés-Hernández, PhD; Narayanaswamy Venketasubramanian, MBBS; Prof Frans R J Verhey, PhD; Prof Joanna M Wardlaw, PhD; Prof David J Werring, PhD; Xu Xin, PhD; Prof Kyung-Ho Yu, PhD; Martine J E van Zandvoort, PhD; Lei Zhao, PhD.

# References


1. Fox, M. D. Mapping Symptoms to Brain Networks with the Human Connectome. *New England Journal of Medicine* **379**, 2237–2245 (2018).
2. Boes, A. D. *et al.* Network localization of neurological symptoms from focal brain lesions. *Brain* **138**, 3061–3075 (2015).
3. van den Heuvel, M. P. *et al.* Investigating the methodological foundation of lesion network mapping. *Nat Neurosci* 1–11 (2026) doi:10.1038/s41593-025-02196-7.
4. Petersen, M. *et al.* Enhancing cognitive performance prediction by white matter hyperintensity connectivity assessment. *Brain* awae315 (2024) doi:10.1093/brain/awae315.
5. Schaper, F. L. W. V. J. *et al.* Mapping Lesion-Related Epilepsy to a Human Brain Network. *JAMA Neurol* **80**, 891 (2023).
6. Klingbeil, J. *et al.* Pathological laughter and crying: insights from lesion network-symptom-mapping. *Brain* **144**, 3264–3276 (2021).
7. Cohen, A. L. *et al.* Looking beyond the face area: lesion network mapping of prosopagnosia. *Brain* **142**, 3975–3990 (2019).
8. Weaver, N. A. *et al.* Strategic infarct locations for post-stroke cognitive impairment: a pooled analysis of individual patient data from 12 acute ischaemic stroke cohorts. *The Lancet Neurology* **0**, (2021).
9. Cohen, A., Al-Fatly, B. & Horn, A. GSP1000 Preprocessed Connectome for Lead DBS. Harvard Dataverse https://doi.org/10.7910/DVN/KKTJQC (2024).
10. Horn, A. & Kühn, A. A. Lead-DBS: A toolbox for deep brain stimulation electrode localizations and visualizations. *NeuroImage* **107**, 127–135 (2015).
11. Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. & Nichols, T. E. Permutation inference for the general linear model. *Neuroimage* **92**, 381–397 (2014).

12. Weaver, N. A. *et al*. The Meta VCI Map consortium for meta-analyses on strategic lesion locations for vascular cognitive impairment using lesion-symptom mapping: Design and multicenter pilot study. *Alzheimers Dement (Amst)* **11**, 310–326 (2019).

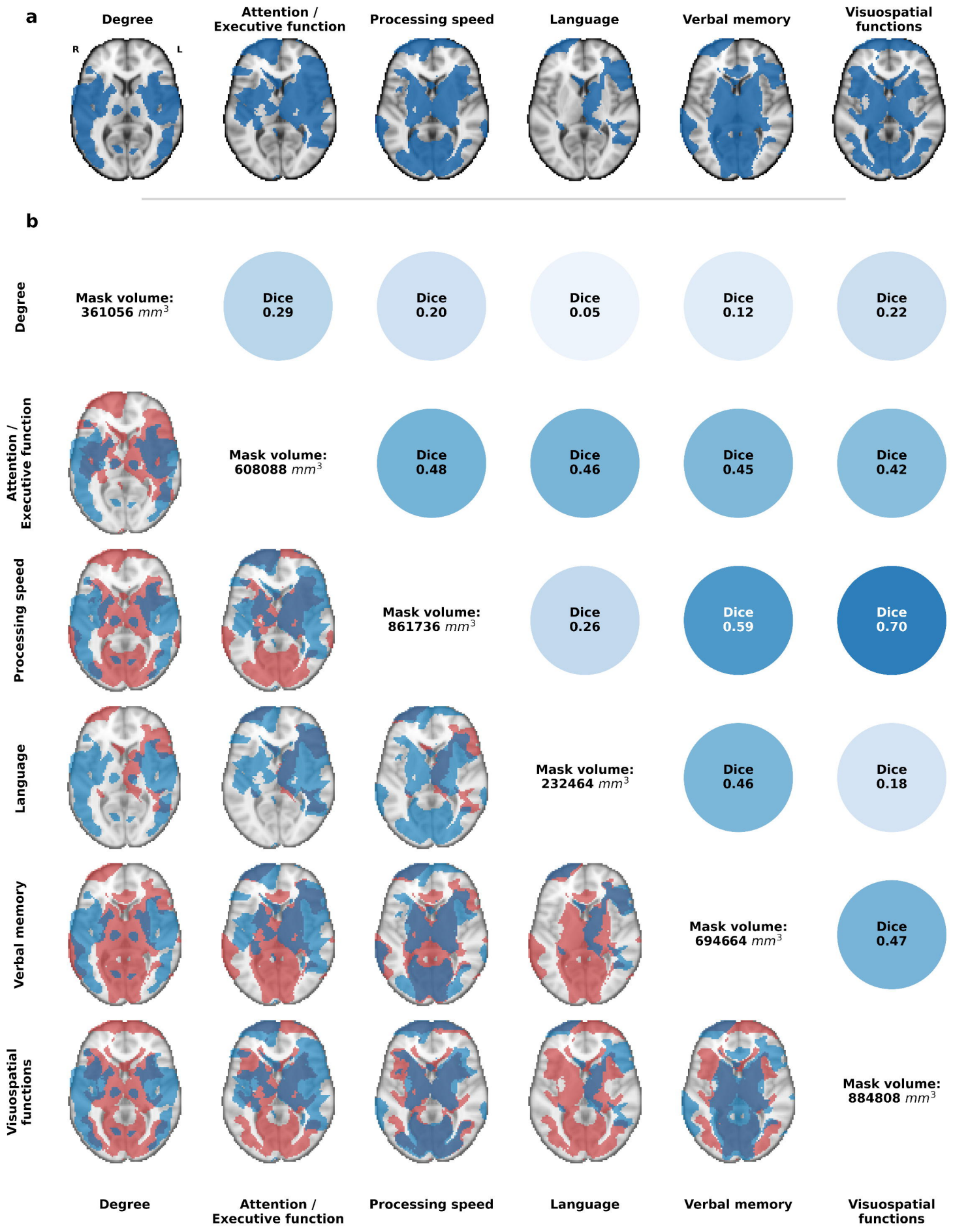
**Competing interests**

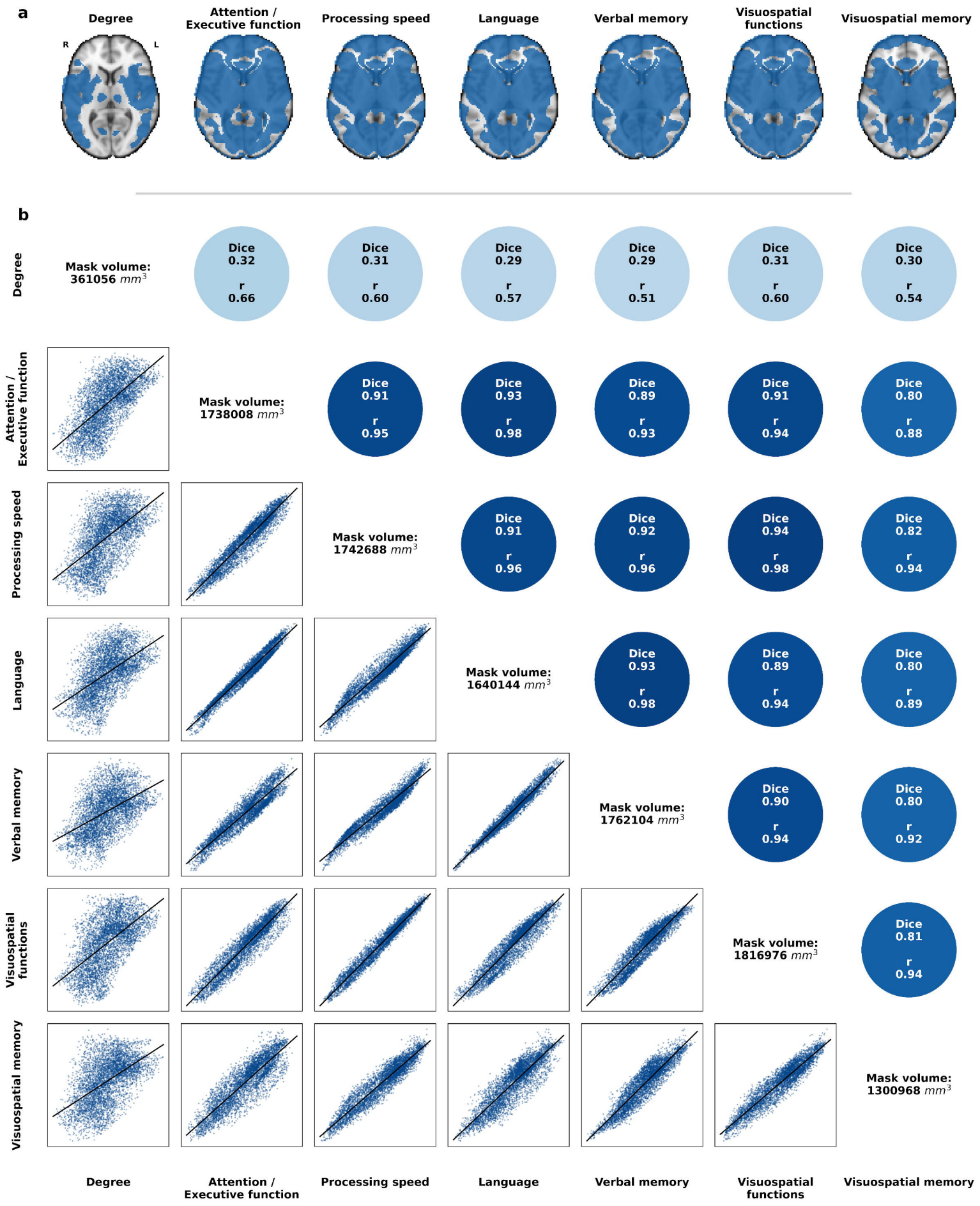
The authors declare no competing interests.

# Figure legends

**Figure 1 | Permutation-based inference reduces convergence toward degree-like structure and decreases cross-domain similarity.** (a) Symptom-linked lesion network maps derived using label permutation, thresholded at $p_{FDR} < 0.05$. (b) Pairwise comparisons of label-permutation lesion network maps. Upper triangle: similarity quantified as Dice overlap of significance-thresholded maps. The first row/column compares each map with the normative connectome degree map (degree defined as the row-sum of the normative FC matrix; binary degree mask thresholded at the 95th percentile). Lower triangle: pairwise overlaps of significance-thresholded maps; maps corresponding to columns are shown in blue and rows in red. Diagonal: volumes of thresholded masks ($mm^3$).

**Figure 2 | Parametric symptom-linked lesion network mapping yields highly similar maps across cognitive domains.** (a) Symptom-linked lesion network maps derived from parametric statistics, thresholded at $p_{FDR} < 0.05$. (b) Pairwise comparisons of symptom-linked maps. Upper triangle: similarity quantified as Dice overlap of thresholded maps and Pearson correlation (r) of voxelwise T-statistics. The first row/column compares each symptom-linked map with the normative connectome degree map (degree defined as the row-sum of the normative FC matrix; binary degree mask thresholded at the 95th percentile). Lower triangle: voxelwise scatterplots of T-statistics. Diagonal: volumes of thresholded masks ($mm^3$).


a
Degree
Attention / Executive function
Processing speed
Language
Verbal memory
Visuospatial functions
R
L
b
Degree
Attention / Executive function
Processing speed
Language
Verbal memory
Visuospatial functions
Mask volume: 361056 mm3
Dice 0.29
Dice 0.20
Dice 0.05
Dice 0.12
Dice 0.22
Mask volume: 608088 mm3
Dice 0.48
Dice 0.46
Dice 0.45
Dice 0.42
Mask volume: 861736 mm3
Dice 0.26
Dice 0.59
Dice 0.70
Mask volume: 232464 mm3
Dice 0.46
Dice 0.18
Mask volume: 694664 mm3
Dice 0.47
Mask volume: 884808 mm3
Degree
Attention / Executive function
Processing speed
Language
Verbal memory
Visuospatial functions


a
Degree
Attention / Executive function
Processing speed
Language
Verbal memory
Visuospatial functions
Visuospatial memory
R
L
b
Degree
Mask volume: 361056 mm3
Dice 0.32 r 0.66
Dice 0.31 r 0.60
Dice 0.29 r 0.57
Dice 0.29 r 0.51
Dice 0.31 r 0.60
Dice 0.30 r 0.54
Attention / Executive function
Mask volume: 1738008 mm3
Dice 0.91 r 0.95
Dice 0.93 r 0.98
Dice 0.89 r 0.93
Dice 0.91 r 0.94
Dice 0.80 r 0.88
Processing speed
Mask volume: 1742688 mm3
Dice 0.91 r 0.96
Dice 0.92 r 0.96
Dice 0.94 r 0.98
Dice 0.82 r 0.94
Language
Mask volume: 1640144 mm3
Dice 0.93 r 0.98
Dice 0.89 r 0.94
Dice 0.80 r 0.89
Verbal memory
Mask volume: 1762104 mm3
Dice 0.90 r 0.94
Dice 0.80 r 0.92
Visuospatial functions
Mask volume: 1816976 mm3
Dice 0.81 r 0.94
Visuospatial memory
Mask volume: 1300968 mm3
Degree
Attention / Executive function
Processing speed
Language
Verbal memory
Visuospatial functions
Visuospatial memory