Revisiting the methodological foundation of 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

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 Jülich, Germany

Correspondence to: m.petersen-28@umcutrecht.nl

Abstract

Lesion network mapping (LNM) links focal brain lesions to distributed neural circuits by projecting lesion locations through a normative functional connectome. van den Heuvel and colleagues recently showed how commonly used LNM procedures generate maps that converge on nonspecific, low-dimensional properties of the connectome, introducing a bias. Consequently, many published maps of different conditions appear strikingly similar. Here, we offer an alternative approach that does highlight distinct symptom-specific signals in LNM. In a multicenter dataset of 2,950 stroke patients, we replicate the expected convergence under the standard procedures, but also demonstrate how permuting symptom labels provides an appropriate null model that delivers distinct, biologically plausible networks for specific cognitive functions.

Main

LNM is an analytical framework that combines lesion and functional connectivity (FC) information to identify brain networks related to neurological and psychiatric conditions. It formalizes the idea that anatomically distinct lesions can produce similar symptoms when they perturb the same functional network.

LNM produces statistical maps highlighting network regions that are associated with a symptom or condition using multiple analysis steps. van den Heuvel et al. alerted the field to remarkable, and biologically implausible, similarity of lesion network maps for very different conditions and function, due to a form of bias that is inherent to analytical approaches that are common in LNM.[1] In their work, van den Heuvel et al. reformulated this standard LNM approach as a simplified linear operation on a fixed normative FC matrix.[1] Participant-level lesion-derived connectivity maps are obtained by multiplying a lesion matrix (M) with a fixed normative FC matrix (C). A sample-level map is then obtained by aggregating participant maps: either via averaging or summation, or – in symptom-linked LNM – via statistical association with symptom scores. Using this reformulation, the authors prove that group-level aggregation across lesions/participants repeatedly samples the same connectome matrix C, biasing resulting aggregate maps toward elementary, nonspecific properties of C – most prominently its row sum (node degree), and other low-rank components. Consequently, markedly different lesion sets can yield highly similar lesion network maps.

We agree that this striking convergence warrants a reappraisal of how many published LNM findings are interpreted: under standard LNM procedures, similarity between maps should

not be taken as evidence of shared neuroanatomical correlates across symptoms or conditions. This analysis underscores the need for more rigorous inference, including appropriate null models and validation. One plausible contributor to the authors' findings is that standard LNM workflows rely on parametric tests, which implicitly assume independence in a setting where maps can strongly converge due to shared connectome structure.[1] We argue that nonparametric permutation-based techniques offer a solution to this problem. Inference based on symptom-label permutation is well established in neuroimaging and has also been adopted in prior LNM studies, including our own.[6–8] However, its role as an appropriate null for addressing connectome-driven convergence has not been examined. The label permutation strategy retains the structured covariance and latent-factor content induced by the normative connectome, i.e., the main source of convergence, but breaks the mapping between symptom and connectome-derived features. Because the convergence-related artifact is preserved, it is reflected in the null distribution; consequently, empirical effects driven mainly by the artifact will not be statistically significant. The resulting null therefore tests whether symptom-linked effects exceed what is expected under exchangeability of labels, given connectome-induced dependence among the maps.

To evaluate standard parametric voxelwise inference against symptom-label permutation in LNM, we analyzed multicenter data of post-stroke cognitive impairment patients from the Meta VCI Map Consortium (n = 2,950, 12 cohorts). We statistically compared cognitively impaired versus unimpaired stroke patients based on binary domain-impairment labels for attention/executive function, processing speed, language, verbal memory, visuospatial function and visuospatial memory.[2] Using normalized acute stroke lesion segmentations, we computed individual lesion connectivity maps employing Lead-DBS with the GSP1000 normative functional connectome.[3,4] We then contrasted two inference approaches: (i) a standard parametric voxelwise model, as also employed in van den Heuvel's paper, and (ii) permutation-based inference, permuting symptom labels relative to the fixed connectivity maps. We quantified the pairwise similarity of resulting domain-level lesion network maps using spatial correlation of T-statistics and Dice overlap of thresholded significance maps ($p_{FDR} < 0.05$). To determine whether the label permutation approach yields information beyond lesion location, we also compared it to traditional voxel-based lesion-symptom mapping. Statistical analyses were implemented using Permutation Analysis of Linear Models (PALM) adjusting for lesion volume and correcting p-values for false-discovery-rate.[5] Permutation-based testing involved 1000 permutations and threshold-free cluster enhancement.

As predicted by van den Heuvel et al., parametric maps corresponded with the normative degree map (mean r = 0.58 ± 0.05; mean Dice = 0.30 ± 0.01) and showed biologically implausible high similarity across symptom cognitive domains (mean r = 0.94 ± 0.03; mean Dice = 0.88 ± 0.05) (***Figure 1***). For example, the parametric language map – where lateralization is expected – showed limited apparent specificity, with substantial alignment to domain-general connectome structure. However, employing label-permutation nulls yielded lesion network maps differing from parametric results (***Figure 2***). Overlap with the degree map was lower and more variable (mean Dice = 0.18 ± 0.08). Cross-domain similarity was also reduced (mean Dice = 0.45 ± 0.14), indicating that maps no longer converged onto a shared, domain-general pattern. In practical terms, the null procedure yielded maps with greater anatomical specificity. This anatomical specificity was especially evident for language. The language network preferentially highlighted regions that align with a priori expectations for language circuitry. Visuospatial memory showed no significant group differences when evaluated against the label permutation null. 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 dependence-preserving inference does not reduce LNM to focal lesion effects. 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.

By confirming the central limitation of standard LNM procedures identified by van den Heuvel et al. and demonstrating a viable alternative, we clarify 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. The effect size of the remaining symptom-linked component is therefore expected to be smaller and harder to detect but more specific. The persistence of nontrivial maps for several domains under this dependence-preserving null supports the presence of residual symptom-linked effects. By contrast, the absence of significant effects for visuospatial memory underscores that not all outcomes remain detectable once stringent control is imposed.

On this basis, we recommend grounding LNM inference in label permutation or other exchangeability-respecting nulls that explicitly preserve connectome-induced dependence. This shift would separate robust effects from shared structure imposed by the normative connectome 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 suggests that a first-principles revision of the technique may be closer at hand than expected.

References

1. 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.

2. 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).

3. Horn, A. & Kühn, A. A. Lead-DBS: A toolbox for deep brain stimulation electrode localizations and visualizations. *NeuroImage* **107**, 127–135 (2015).

4. Cohen, A., Al-Fatly, B. & Horn, A. GSP1000 Preprocessed Connectome for Lead DBS. Harvard Dataverse https://doi.org/10.7910/DVN/KKTJQC (2024).

5. 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).

6. Cohen, A. L. *et al.* Looking beyond the face area: lesion network mapping of prosopagnosia. *Brain* **142**, 3975–3990 (2019).

7. Petersen, M. *et al.* Enhancing cognitive performance prediction by white matter hyperintensity connectivity assessment. *Brain* awae315 (2024) doi:10.1093/brain/awae315.

8. Schaper, F. L. W. V. J. *et al.* Mapping Lesion-Related Epilepsy to a Human Brain Network. *JAMA Neurol* **80**, 891 (2023).

Competing interests

The authors declare 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 acknowledged in the supplementary materials.

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.

Code availability

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.

Figures

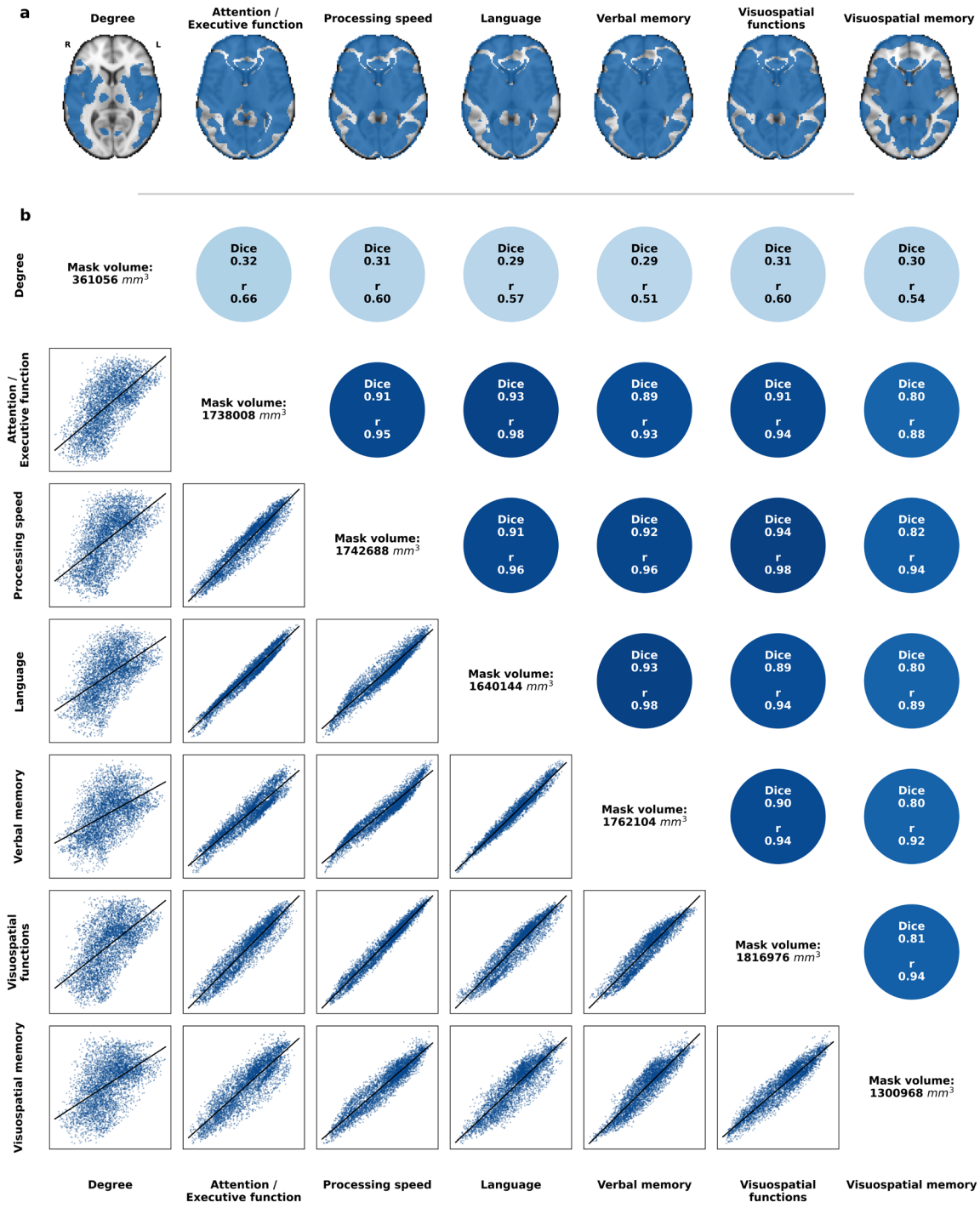


Figure 1 | Parametric symptom-linked lesion network mapping yields highly similar maps across cognitive domains. (a) Symptom-linked lesion network maps derived from parametric statistics, thresholded at $p_{FDR} < 0.05$. (b) Pairwise comparisons of symptom-linked maps. Upper triangle: similarity quantified 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).

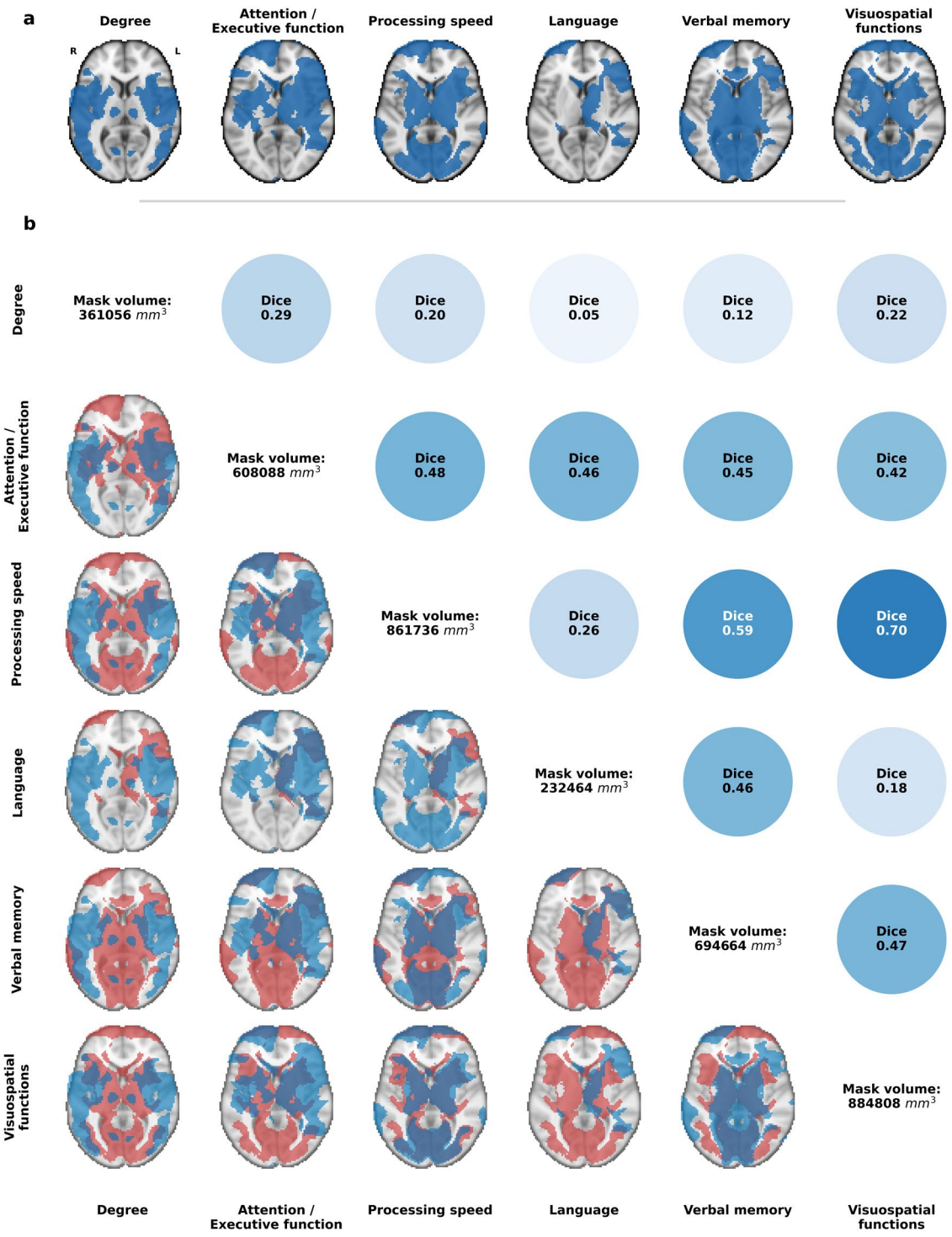
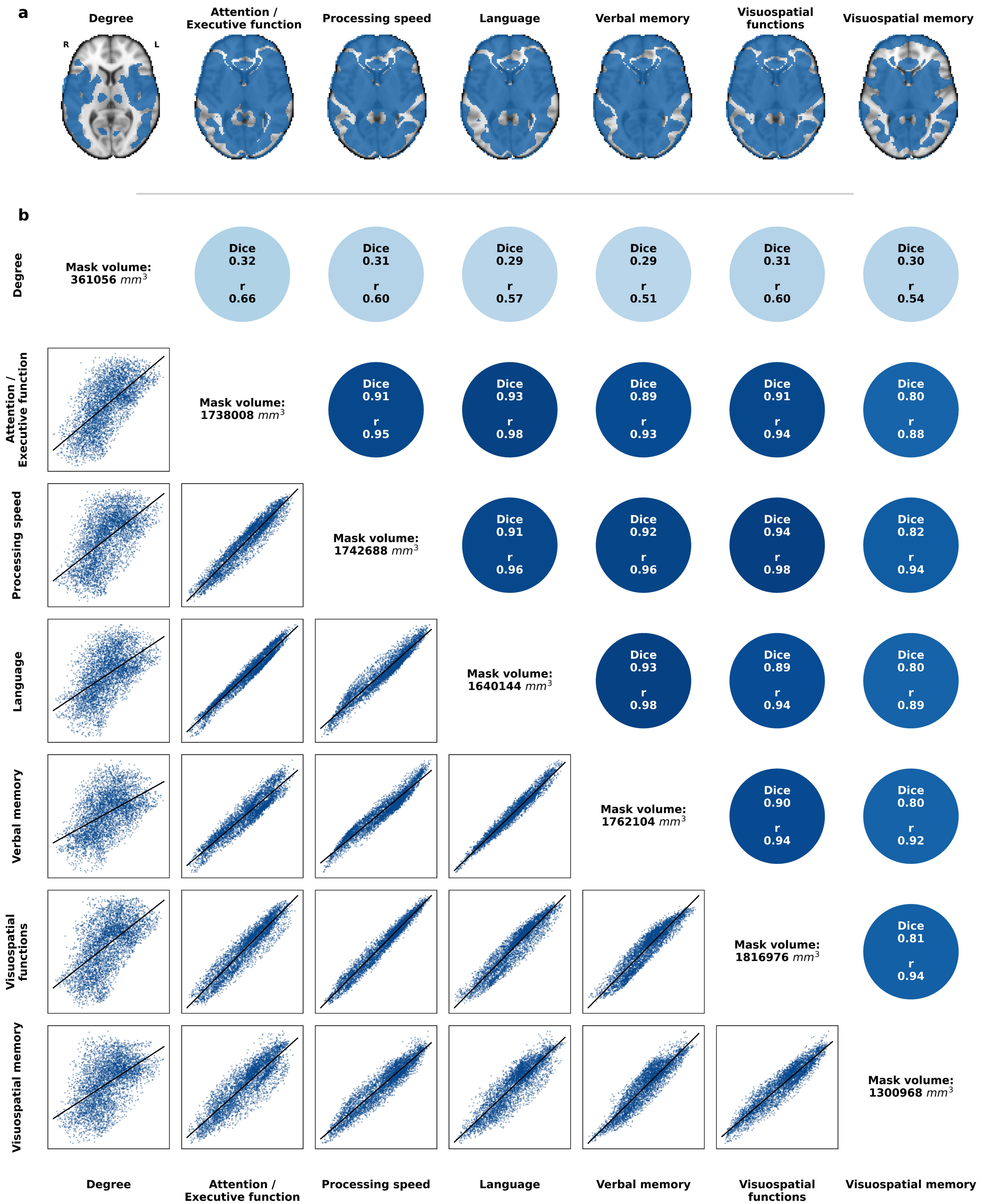
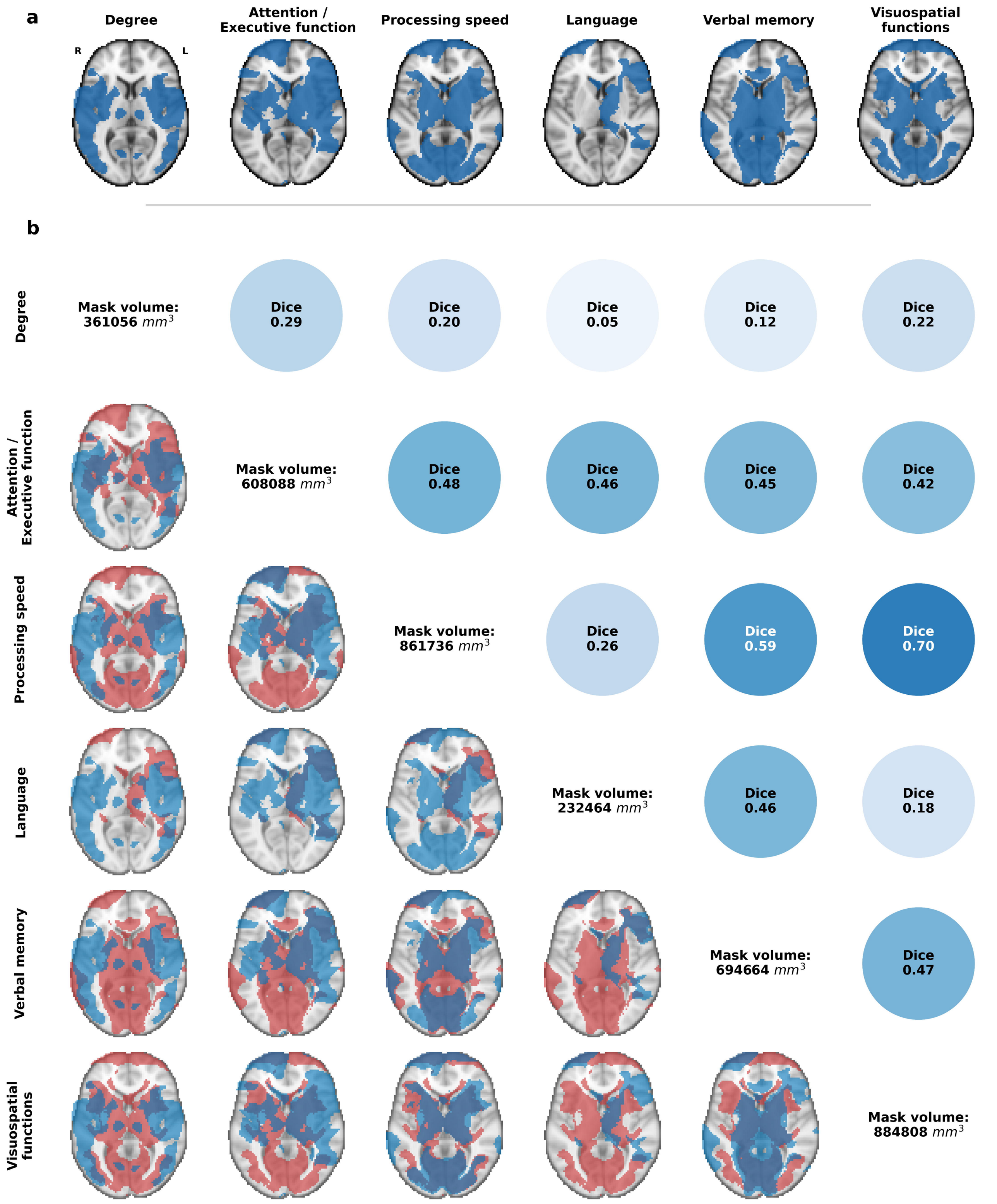


Figure 2 | Permutation-based inference reduces convergence toward degree-like structure and decreases cross-domain similarity. (a) Symptom-linked lesion network maps derived using label permutation, thresholded at $p_{FDR} < 0.05$. (b) Pairwise comparisons of label-permutation lesion network maps. Upper 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 (mm³).


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


a
Degree
Attention / Executive function
Processing speed
Language
Verbal memory
Visuospatial functions
R
L
b
Degree
Mask volume: 361056 mm^3
Dice 0.29
Dice 0.20
Dice 0.05
Dice 0.12
Dice 0.22
Attention / Executive function
Mask volume: 608088 mm^3
Dice 0.48
Dice 0.46
Dice 0.45
Dice 0.42
Processing speed
Mask volume: 861736 mm^3
Dice 0.26
Dice 0.59
Dice 0.70
Language
Mask volume: 232464 mm^3
Dice 0.46
Dice 0.18
Verbal memory
Mask volume: 694664 mm^3
Dice 0.47
Visuospatial functions
Mask volume: 884808 mm^3
Degree
Attention / Executive function
Processing speed
Language
Verbal memory
Visuospatial functions