Title

To smooth or not to smooth: One step closer to single-voxel accuracy without spatial smoothing

Authors

Eileen Luders[1,2,3], Robert Dahnke[4], Christian Gaser[4]*, Alzheimer's Disease Neuroimaging Initiative†

Affiliations

[1]School of Psychology, University of Auckland, Auckland, New Zealand
[2]Laboratory of Neuro Imaging, School of Medicine, University of Southern California, Los Angeles, USA
[3]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden
[4]Departments of Psychiatry and Neurology, Jena University Hospital, Jena, Germany

†Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

*Correspondence should be addressed to:
Christian Gaser, Ph.D. (christian.gaser@uni-jena.de)

Number of Pages: **13**
Number of Figures: **2**

Keywords: **spatial smoothing, TFCE, neuroscience, non-parametric statistics, MRI, human brain mapping**

Abstract

Traditionally, when conducting voxel- or vertex-wise analyses in neuroimaging studies, it seemed imperative that brain data are convoluted with a Gaussian kernel – a procedure known as "spatial smoothing". However, we argue that smoothing may be omitted entirely for the benefit of a dramatically enhanced regional specificity under certain conditions. We demonstrate the suitability of this omission by combining high-dimensional warping and threshold-free cluster enhancement (TFCE) in a sample of 754 brains. Our findings suggest that without the traditional smoothing step it is possible to dissociate neighboring brain areas with an accuracy of a single voxel, altogether taking the field of human brain mapping to a new level. Nevertheless, classic analyses based on smoothed data will continue to provide important insights, especially if parametric tests are required, if image data cannot be precisely aligned across individuals or time points, and/or if anticipated effect sizes are only small.

Introduction

Spatial smoothing is a key preprocessing step in the field of brain mapping, where image data are convoluted with a (Gaussian) kernel before they are statistically analyzed. Smoothing ensures that data are normally distributed (a pre-requisite for conducting parametric tests), accounts for any remaining anatomical variations across brain scans after the spatial registration, and improves the signal-to-noise ratio (which increases the sensitivity of the statistical analysis). However, smoothing ultimately also poses a problem as it severely impacts the regional specificity of the analysis outcomes (the larger the smoothing kernel, the lower the specificity). As a result, when generating the statistical map, significance clusters often spread across anatomical boundaries (the larger the smoothing kernel, the higher the spread). That is, valuable regional information originally inherent in the acquired brain image gets lost during the smoothing procedure and spatial accuracy decreases.

In other words, there are considerable drawbacks to smoothing and the question arises whether we can do without spatial smoothing. We strongly believe an omission of the smoothing step is a viable option, as long as the following conditions are met: (I) The applied statistical approach does not require a normal distribution of data; (II) brain images are close to perfectly aligned; (III) and statistical degrees of freedom and/or anticipated effect sizes are sufficiently large. Ensuring that all three conditions are met may seem challenging, but is possible considering new developments in the field of neuroscience and human brain mapping. With respect to condition I, there are numerous statistical approaches based on non-parametric inference (i.e., data do not need to be normally distributed), among them threshold-free cluster enhancement (TFCE; *1*). TFCE has the additional benefit of being relatively sensitive (which is also relevant for condition 3), achieved by integrating cluster information (cluster size significance) with voxel-wise statistical inference (peak voxel significance). With respect to

condition II, there are a multitude of spatial registration algorithms (*2*) that have been substantially improved over the years allowing nowadays for an almost perfect overlap and spatial correspondence across brain images, among them high-dimensional warping (*2-4*). With respect to condition III, there is an ever increasing pool of large-scale databases, often containing hundreds or even thousands of brain scans, among them ADNI (https://adni.loni.usc.edu/) or the UK biobank (https://www.ukbiobank.ac.uk/). Compiling samples with higher numbers of participants / brains results in higher degrees of freedom and as such an appropriate sensibility of the statistical analysis. If larger sample sizes are not an option, naturally the detectability of effects would also be increased when expecting big effect sizes, such as in clinical conditions affecting the brain (e.g., neurodegenerative diseases).

Here, we leveraged a large dataset (n=754) from the Alzheimer's Disease Neuroimaging Initiative (ADNI), to test whether analyzing unsmoothed data affords an adequate sensitivity and leads to an improved regional specificity compared to analyzing smoothed data using standard kernel sizes of 2 mm, 4 mm, and 6 mm full width at half maximum (FWHM). For this purpose, we conducted a voxel-based morphometry (VBM) analysis, in association with TFCE (*1*) and high-dimensional warping (*4*), comparing voxel-wise gray matter between four subgroups: healthy controls (CTL; n=218), individuals with stable Mild Cognitive Impairment (sMCI; n=222), individuals with progressive MCI (pMCI; n=130), and individuals with Alzheimer's disease (AD; n=184).

Results

As shown in **Figure 1**, the most pronounced effects (see red clusters) were detected bilaterally in the hippocampus, amygdala, as well as the parahippocampal and entorhinal gyrus, with significantly more gray matter in CTL than in sMCI, pMCI, and AD. While these effects were

evident in both unsmoothed and smoothed data, findings were spread wider and clusters bled over anatomical boundaries in smoothed data (the larger the FWHM, the larger the spread).

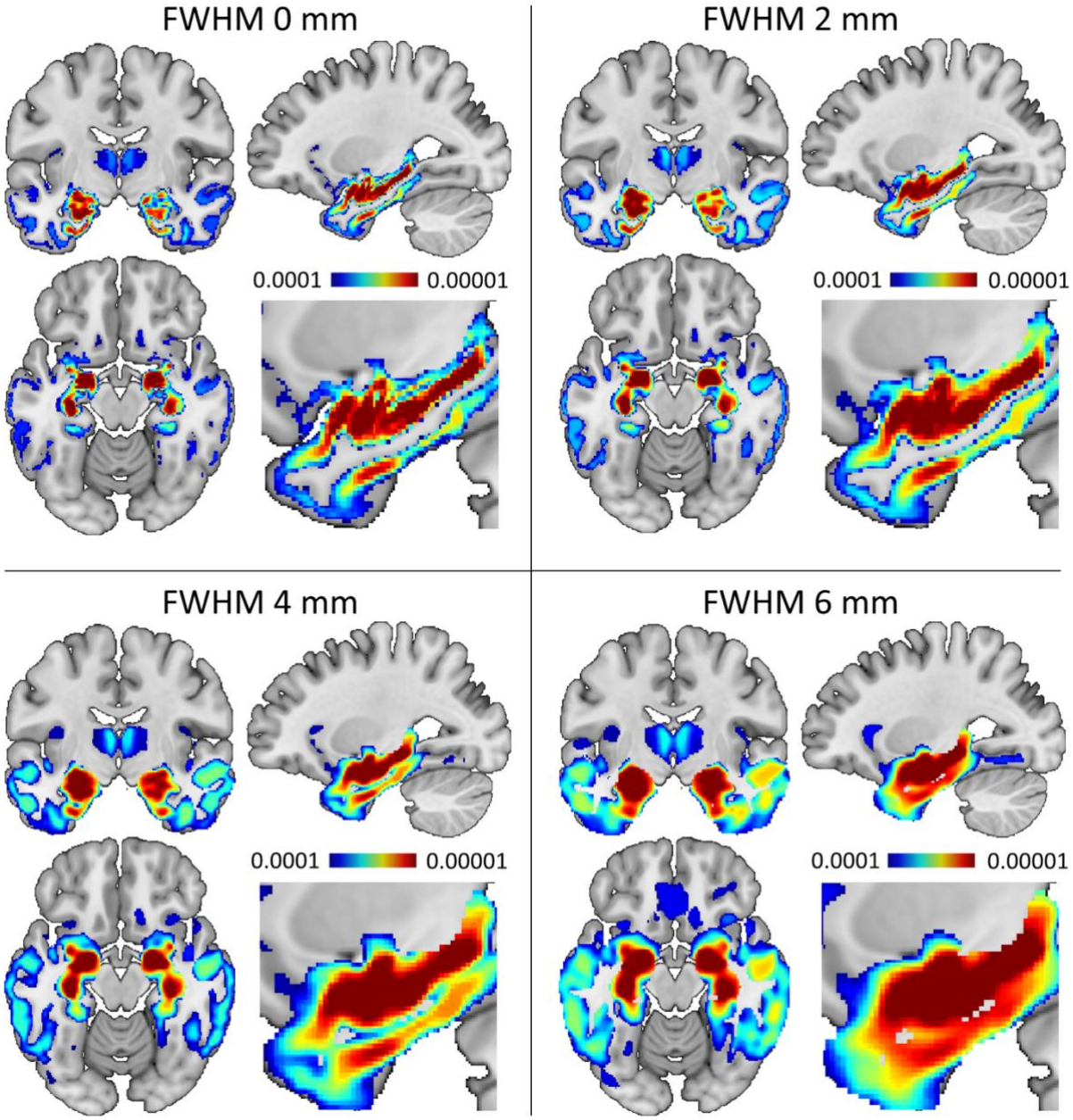

Fig. 1. Significance maps across the four datasets. All outcomes were derived for contrast CTL>sMCI>pMCI>A at p<0.0001 using TFCE and family-wise error (FWE). Unsmoothed data (FWHM = 0 mm) and data smoothed with different kernel sizes (FWHM = 2 mm, 4 mm, and 6 mm. Orthogonal slices at x; y; z = -28 mm; -11 mm; -15 mm. The extra sagittal slice (x = -28) provides a close-up of the hippocampal complex. The color bar encodes significance (p) ranging between 0.0001 (blue) and 0.00001 (red).

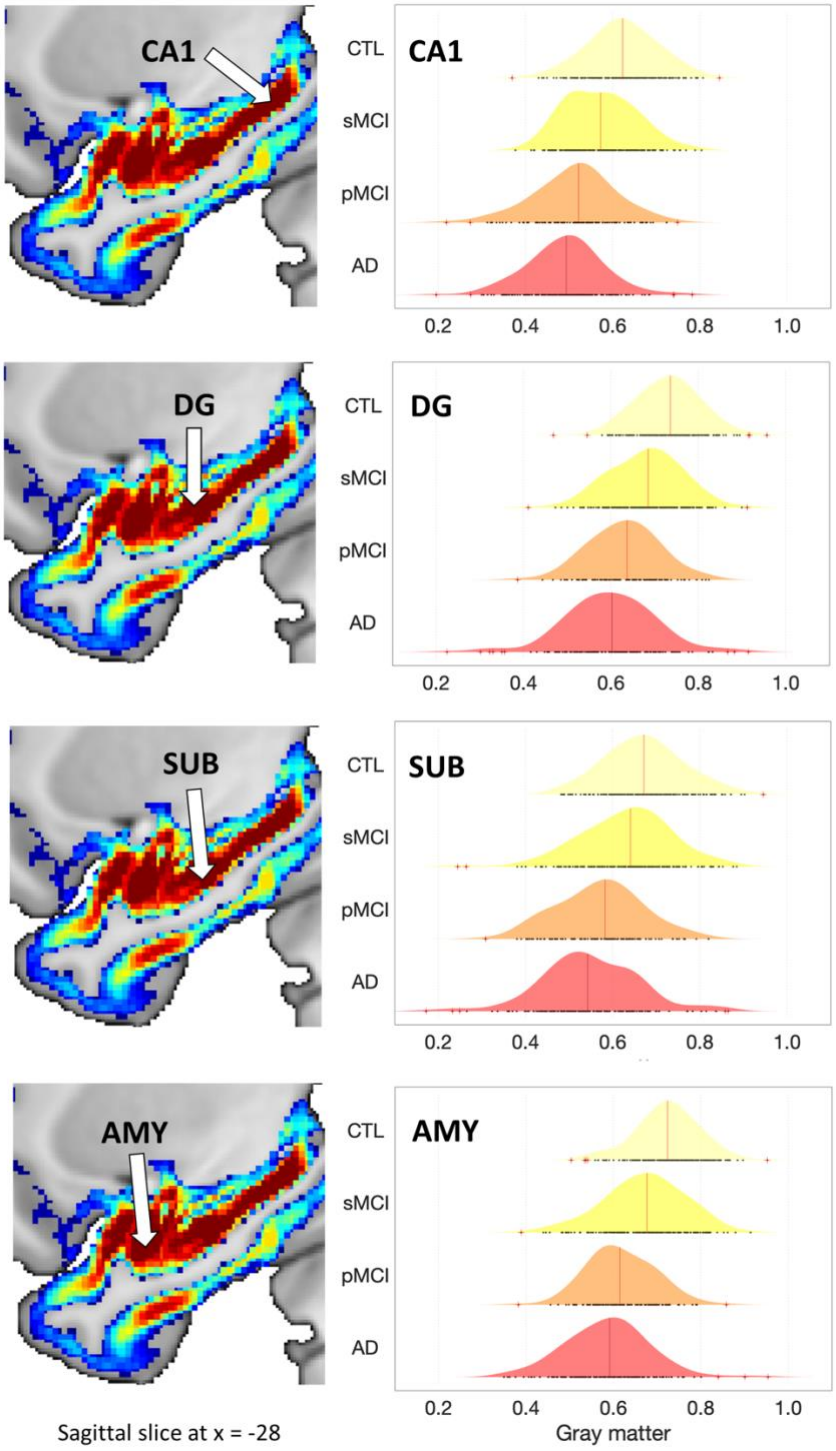

Fig. 2. Density plots for regions of the hippocampal complex in unsmoothed data (FWHM = 0 mm). Right Panel: The x-axes show the amount of voxel-wise gray matter, the y-axes the four different groups: HC = healthy controls, sMCI = stable mild cognitive impairment, pMCI = progressive mild cognitive impairment, and AD = Alzheimer's disease. Left Panel: The arrows point to the approximate voxel for which the gray matter was plotted: CA1 = cornu ammonis (x; y; z = -28 -33 -11), DG = dentate gyrys (x; y; z = -28 -15 -20), SUB = subiculum (x; y; z = -28 -15 -25), and AMY = amygdala (x; y; z = -28 -6 -19). The significance map is identical with the one provided in Figure 1 (i.e., CTL>sMCI>pMCI>AD at $p<0.0001$ using TFCE and FWE).

In contrast, as further depicted in **Figure 2**, in unsmoothed data, the statistical map preserved the regional information of the initial brain scan. It is even possible to identify and discriminate between significance clusters pertaining to known subfields of the hippocampal complex and adjacent regions, such as the cornu ammonis, the dentate gyrus, the subiculum, and the amygdala.

Discussion

There are usually more than a million voxels in a high-resolution brain scan. Thus, any statistical procedures require proper corrections for multiple comparisons. Random Field Theory (RFT), which accounts for spatial dependency in the data (*5*) has been one of the most commonly used correction-method in the field of human brain mapping. However, RFT-based corrections are only valid when image data are normally distributed and as such do not violate the assumption of parametric testing (*6*). As outlined above, spatial smoothing is a means to ensure that data are normally distributed, but spatial smoothing also has the negative side effect of decreasing the regional specificity of the findings. While smoothing on the surface is considered substantially less deleterious than in the volume (*7, 8*), high amounts of smoothing on the surface will still degrade the spatial localization of a cortical area.

Here, we demonstrate that, smoothing may be omitted entirely for the benefit of a dramatically enhanced regional specificity. For this purpose, we analyzed both smoothed and unsmoothed data. In either case, we abstained from parametric statistics altogether and, instead, used a non-parametric TFCE approach. In addition, we applied high-dimensional warping ensuring that all brain images were precisely aligned and in spatial correspondence with each other. We also leveraged a relatively big sample (N=754), containing healthy controls as

well as participants with MCI and AD, altogether increasing the detectability of any effects. While significant group differences were evident in both smoothed and unsmoothed data, effects in unsmoothed data were less regionally specific with clusters bleeding over anatomical boundaries. It is beyond the scope of this methodically driven article to provide an in-depth interpretation of the findings, but it should be at least briefly pointed out that the most pronounced effects (i.e., significantly more gray matter in CTL than in AD) were detected in the hippocampal complex and amygdala, which is well in line with other reports in the literature (*9-11*).

Skipping the spatial smoothing step may seem radical, but evidence for the suitability of this approach is increasing (*11*), especially if brain images are precisely aligned (*12, 13*). In fact, there seems to be a shift in attitude towards smoothing in the field of human brain mapping, complementing traditional views that spatial smoothing is absolute necessary with more differentiated views (*8, 11, 14-16*). The current findings are in close resemblance with this shift in attitude. Omitting the smoothing step made it possible to dissociate neighboring brain areas with a single-voxel accuracy. Nevertheless, this gain in regional specificity was set off by a slight loss of sensitivity, which is illustrated by the more pronounced significance in the extended clusters in the smoothed data. The effect size in the unsmoothed data was still big enough to pass the threshold of significance, so this did not constitute a problem.

In summary, based on the outcomes of the current analysis, we conclude that TFCE without smoothing leads to an improved regional specificity and that its trade-offs in sensitive warrant its application over standard parametric tests with smoothing. However, the decreased sensitivity might pose a problem in other analyses with thresholds at the border of significance. Therefore, analyses based on smoothed data will continue to provide important insights,

especially if parametric tests are required, if anticipated effect sizes are only small, and/or if image data cannot be precisely aligned.

Materials and Methods

Study Sample

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). More specifically, we used the T1-weighted MRI data from ADNI 1, where baseline MRI data and test scores in selected cognitive scales (i.e., Mini-Mental State Examination [MMSE]) were available. Altogether, the sample contained 754 individuals, who were classified into four groups as (I) healthy controls (CTL) if they were cognitively healthy at baseline as well as at three-year follow-up (n = 218; males/females = 111/107, mean/SD age = 76.03/5.04 years, mean/SD MMSE = 29.13/1.01); (II) individuals with stable MCI (sMCI), if they were diagnosed with MCI at baseline as well as at three-year follow-up (n = 222, males/females = 145/77, mean/SD age = 75.51/7.32 years, mean/SD MMSE = 27.20/1.79), (III) individuals with progressive MCI (pMCI) if they were diagnosed with MCI at baseline and with AD at some point during the three-year follow-up without reversion to MCI (n = 130, males/females = 79/51, mean/SD age = 74.66/7.12 years, mean/SD MMSE = 26.68/1.75), and (IV) individuals with AD (AD), if they were diagnosed with AD at baseline as well as at three-year follow-up (n = 184, males/females = 95/99, mean/SD age = 75.28/7.56 years, mean/SD MMSE = 23.25/2.04).

Data Processing

All brain images were processed using the CAT12 toolbox (r1940; http://www.neuro.uni-jena.de/cat), as implemented in SPM12 (https://www.fil.ion.ucl.ac.uk/spm/software/spm12/). The default CAT12 settings were applied for bias field correction (*17*) and tissue classification, which was based on adaptive maximum a posteriori estimations (*18*) and also accounted for partial volume effects (*19*). In contrast, for the spatial registration, we used a detailed geodesic shooting template with a spatial resolution of 1 mm, and also saved the spatially registered gray matter segments with a spatial resolution of 1 mm (the default is 1.5 mm). The resulting gray matter segments underwent visual and automated quality checks and were finally multiplied generating four identical sets of data: The first set remained unsmoothed; the other three sets were smoothed with a Gaussian kernel of 2, 4, and 6 mm FWHM. To mitigate any segmentation artifacts on the gray/white border, an absolute gray matter threshold of 0.1 was applied to the unsmoothed data; the same mask was also applied to the other three sets (to avoid any bias due to different masks).

Statistical Analysis

For each of the four datasets (0, 2, 4, and 6 mm FWHM, respectively), we ran an ANCOVA with four groups (CTL, sMCI, pMCI, and AD), while removing the variance associated with total intracranial volume. The latter was calculated by adding the volumes of the gray matter, white matter, and cerebrospinal fluid segments in native space. To test for increasing atrophy across the four groups (CTL>sMCI> pMCI>AD), the following contrast was applied: 1.5 0.5 -0.5 -1.5. The statistical analysis was conducted using the non-parametric TFCE-toolbox (which is freely available at http://www.neuro.uni-jena.de/tfce) running 25,000 permutations and applying a threshold of 0.0001 that was family-wise error (FWE) corrected.

Acknowledgments
Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Funding:
Carl Zeiss Foundation
Alexander von Humboldt Foundation
Marie Skłodowska-Curie Innovative Training Network (859890-SmartAge-H2020-MSCA-ITN-2019)

Author contributions: Conceptualization: CG; Methodology: RD, CG; Investigation: CG; Visualization: EL, RD, CG; Writing—original draft: EL; Writing—review & editing: EL, RD, CG

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: The sources for all data and analyses are provided in the main text.

References

1. S. M. Smith, T. E. Nichols, Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* **44**, 83-98 (2009).
2. A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M. C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, R. V. Parsey, Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage* **46**, 786-802 (2009).
3. B. B. Avants, C. L. Epstein, M. Grossman, J. C. Gee, Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* **12**, 26-41 (2008).
4. J. Ashburner, K. J. Friston, Diffeomorphic registration using geodesic shooting and Gauss-Newton optimisation. *Neuroimage* **55**, 954-967 (2011).
5. K. J. Worsley, A. C. Evans, S. Marrett, P. Neelin, A three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab* **12**, 900-918 (1992).
6. S. Hayasaka, K. L. Phan, I. Liberzon, K. J. Worsley, T. E. Nichols, Nonstationary cluster-size inference with random field and permutation methods. *NeuroImage* **22**, 676-687 (2004).
7. S. Brodoehl, C. Gaser, R. Dahnke, O. W. Witte, C. M. Klingner, Surface-based analysis increases the specificity of cortical activation patterns and connectivity results. *Scientific reports* **10**, 5737 (2020).
8. T. S. Coalson, D. C. Van Essen, M. F. Glasser, The impact of traditional neuroimaging methods on the spatial localization of cortical areas. *Proc Natl Acad Sci U S A* **115**, E6356-E6365 (2018).
9. M. I. Miller, L. Younes, J. T. Ratnanather, T. Brown, T. Reigel, H. Trinh, X. Tang, P. Barker, S. Mori, M. Albert, Amygdala Atrophy in MCI/Alzheimer's Disease in the BIOCARD cohort based on Diffeomorphic Morphometry. *Med Image Comput Comput Assist Interv* **2012**, 155-166 (2012).
10. S. P. Poulin, R. Dautoff, J. C. Morris, L. F. Barrett, B. C. Dickerson, Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity. *Psychiatry Res* **194**, 7-13 (2011).
11. R. Turner, S. Geyer, Comparing like with like: the power of knowing where you are. *Brain Connect* **4**, 547-557 (2014).
12. E. C. Robinson, K. Garcia, M. F. Glasser, Z. Chen, T. S. Coalson, A. Makropoulos, J. Bozek, R. Wright, A. Schuh, M. Webster, J. Hutter, A. Price, L. Cordero Grande, E. Hughes, N. Tusor, P. V. Bayly, D. C. Van Essen, S. M. Smith, A. D. Edwards, J. Hajnal, M. Jenkinson, B. Glocker, D. Rueckert, Multimodal surface matching with higher-order smoothness constraints. *NeuroImage* **167**, 453-465 (2018).
13. E. C. Robinson, S. Jbabdi, M. F. Glasser, J. Andersson, G. C. Burgess, M. P. Harms, S. M. Smith, D. C. Van Essen, M. Jenkinson, MSM: a new flexible framework for Multimodal Surface Matching. *NeuroImage* **100**, 414-426 (2014).
14. M. F. Glasser, T. S. Coalson, J. D. Bijsterbosch, S. J. Harrison, M. P. Harms, A. Anticevic, D. C. Van Essen, S. M. Smith, Using temporal ICA to selectively remove global noise while preserving global signal in functional MRI data. *NeuroImage* **181**, 692-717 (2018).

15. M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, S. M. Smith, D. C. Van Essen, A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171-178 (2016).
16. M. F. Glasser, S. M. Smith, D. S. Marcus, J. L. Andersson, E. J. Auerbach, T. E. Behrens, T. S. Coalson, M. P. Harms, M. Jenkinson, S. Moeller, E. C. Robinson, S. N. Sotiropoulos, J. Xu, E. Yacoub, K. Ugurbil, D. C. Van Essen, The Human Connectome Project's neuroimaging approach. *Nature neuroscience* **19**, 1175-1187 (2016).
17. J. Ashburner, K. J. Friston, Unified segmentation. *Neuroimage* **26**, 839-851 (2005).
18. J. C. Rajapakse, J. N. Giedd, J. L. Rapoport, Statistical approach to segmentation of single-channel cerebral MR images. *IEEE Trans.Med.Imaging* **16**, 176-186 (1997).
19. J. Tohka, A. Zijdenbos, A. Evans, Fast and robust parameter estimation for statistical partial volume models in brain MRI. *Neuroimage.* **23**, 84-97 (2004).