bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Joint analysis of area and thickness of the cerebral cortex replaces cortical volume measurements

Anderson M. Winkler[a,b,*], Douglas N. Greve[c], Knut J. Bjuland[d], Thomas E. Nichols[e], Mert R. Sabuncu[c], Asta K. Håberg[f,g], Jon Skranes[d,h], Lars M. Rimol[g]

[a] *Oxford Centre for Functional MRI of the Brain, University of Oxford, OX3 9DU, UK.*
[b] *Department of Psychiatry, Yale University School of Medicine, New Haven, CT, 06511, USA.*
[c] *Martinos Center for Biomedical Imaging, Massachusetts General Hospital/Harvard Medical School, Charlestown, MA, 02129, USA.*
[d] *Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, 7030, Norway.*
[e] *Department of Statistics & Warwick Manufacturing Group, University of Warwick, Coventry, CV4 7AL, UK.*
[f] *Department of Neuroscience, St. Olavs Hospital, Trondheim University Hospital, Trondheim, 7030, Norway.*
[g] *Department of Radiology, St. Olavs Hospital, Trondheim University Hospital, Trondheim, 7030, Norway.*
[h] *Department of Pediatrics, Sørlandet Hospital, 4838, Arendal, Norway.*

Abstract

A detailed investigation of the surface-based methods for assessment of cortical volume and area from magnetic resonance images shows that volume mostly mirrors variability in surface area, while having little sensitivity to cortical thickness, and this remains the case even when volume is assessed using an improved analytic method. Using data from young adults born with very low birth weight and coetaneous controls, it is demonstrated that the permutation-based non-parametric combination (NPC) of thickness and area is more sensitive than volume for studying joint effects on these two quantities, giving equal weight to variation in both. NPC can replace current analyses of cortical volume altogether, allowing a better characterisation of processes that can concomitantly influence brain anatomy, even when their effects are in opposite directions. The method is remarkably general and can be extended to other indices of brain structure and function.

Keywords: Non-Parametric Combination, Brain cortical thickness, Brain cortical area, Brain cortical volume, Areal interpolation

*Corresponding author.
Email address: winkler@fmrib.ox.ac.uk (Anderson M. Winkler)
URL: http://www.fmrib.ox.ac.uk (Anderson M. Winkler)

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1. Introduction

It has been suggested that biological processes that drive horizontal (tangential) and vertical (radial) development of the cerebral cortex are separate from each other (Rakic, 1988; Geschwind and Rakic, 2013), influencing cortical area and thickness independently. These two indices of cerebral morphology are uncorrelated genetically (Panizzon et al., 2009; Winkler et al., 2010), are each influenced by regionally distinct genetic factors (Schmitt et al., 2008; Rimol et al., 2010b; Chen et al., 2012), follow different trajectories over the lifespan (O'Leary et al., 2007; Hogstrom et al., 2013; Fjell et al., 2015), and are differentially associated with cognitive abilities (Schnack et al., 2015; Noble et al., 2015; Vuoksimaa et al., 2016). Moreover, it is cortical area, not thickness, that differs substantially across species (Rakic, 1995). These findings give prominence to the use of surface area alongside thickness in studies of cortical morphology and its relationship to function. The literature contains a variety of approaches and terminologies for its assessment and cortical volume, which commingles thickness and area, continues to be a popular metric, thanks largely to the wide availability of voxel-based morphometry (VBM) (Ashburner and Friston, 2000; Good et al., 2001; Douaud et al., 2007). While analysing cortical thickness and cortical area separately improves specificity (Rimol et al., 2012), it may still be of interest to combine these two measurements so as to increase power to investigate non-specific challenges that could affect thickness and area simultaneously, such as auto-immune (Ceccarelli et al., 2008; Zhang et al., 2016) and infectious neurological disorders (Gitelman et al., 2001; Küper et al., 2011), or in relation to inflammatory markers (Marsland et al., 2008; Zhang et al., 2015). The contributions of this paper are threefold: we (I) expose certain aspects of the various methods for cortical area analysis, in particular the interpolation between surfaces at different resolutions; (II) propose an improved, analytic measurement of volume; (III) show that a joint analysis using the non-parametric combination NPC (Pesarin and Salmaso, 2010; Winkler et al., 2016b) of thickness and area provides a sensible solution to the investigation of factors that can affect cortical morphology, which can replace the analysis of cortical volume altogether.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1.1. Cortical surface area

Comparisons of area across subjects at every point in the cortex depend on registration of the cortical surface and interpolation to a common resolution. Such interpolation must preserve the amount of area at local, regional and global scales, i.e., it must be mass-conservative. A well known interpolation method is the *nearest-neighbour*, which can be enhanced by correction for stretches and shrinkages of the surface during the registration, as available in the function "mris_preproc", part of the FreeSurfer (FS) software package [1]. Another approach is the *retessellation* of the mesh of each subject to the geometry of a common grid, as proposed by Saad et al. (2004) as a way to produce meshes with similar geometry across subjects, thus allowing statistical analysis. Even though the method has been mostly used to compute areal expansion, it can be used for surface area itself, as well as for other areal quantities. A third approach is the use of the barycentric coordinates of each vertex with reference to the vertices of the common grid to *redistribute* the areal quantities, in an approximately mass preserving process. Lastly, a complete framework for analysis of areal quantities was presented in Winkler et al. (2012) using a *pycnophylactic* interpolation method, that is, a method that preserves areal quantities. Table 1 shows an overview of these methods, and a detailed description is in the Experimental Procedures section at the end of the paper.

1.2. Measuring volume and other areal quantities

The volume of cortical grey matter is also an areal quantity, therefore requiring mass-conservative interpolation methods. Volume can be estimated through the use of voxelwise partial volume effects using volume-based representations of the brain, such as in VBM, or from a surface representation, in which it can be measured as the amount of tissue present between the surface placed at the site of the pia-mater, and the surface at the interface between gray and white matter. If the area of either of these surfaces is known, or if the area of a mid-surface, i.e., the surface running half-distance between pial and white surfaces (van Essen, 2005) is known, an estimate of

[1] http://freesurfer.net

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Table 1: Overview of the four different methods to interpolate surface area and areal quantities. A detailed description is in the Experimental Procedures section.

Method	**Description**
Nearest-neighbour	Nearest-neighbour interpolation of areal quantities on the sphere, followed by Jacobian correction.
Retessellation	Barycentric interpolation on the sphere of the native vertex coordinates.
Redistributive	Vertexwise redistribution of areal quantities based on barycentric coordinates of the source in relation to the target.
Pycnophylactic	Mass-conservative facewise interpolation method that uses the overlapping areas between faces of source and target.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

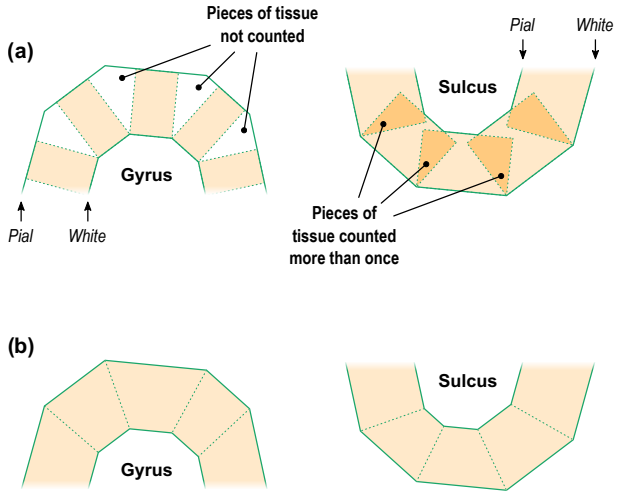


Figure 1: A diagram in two dimensions of the problem of measuring the cortical volume. (*a*) If volume is computed using multiplication of thickness by area, considerable amount of tissue is left unmeasured in the gyri, or measured repeatedly in sulci. The problem is minimised, but not solved, with the use of the mid-surface. (*b*) Instead, vertex coordinates can be used to compute analytically the volume of tissue between matching faces of white and pial surfaces, leaving no tissue under- or over-represented.

the volume can be obtained by multiplying, at each vertex, area by thickness. This procedure, while providing a reasonable approximation that improves over voxel-based measurements, since it is less susceptible to various artefacts (for a discussion of artefacts in VBM, see Ashburner, 2009), is still problematic as it underestimates the volume of tissue that is external to the convexity of the surface, and overestimates volume that is internal to it; both cases are undesirable, and cannot be solved by merely resorting to using an intermediate surface as the mid-surface (Fig.1a). Here a different approach is proposed: each face of the white surface and its matching face in the pial surface are used to define an oblique truncated pyramid, the volume of which is computed analytically, without introducing additional error other than what is intrinsic to the placement of these surfaces (Fig.1b for a 2-D schema and Fig.2 for 3-D).

Quantitative measurements, such as from positron emission tomography (PET), cerebral blood flow, cerebral blood volume, the mass, or number of molecules of a given compound (Leahy and Qi, 2000; van den Hoff, 2005), are all areal quantities whenever these are expressed in absolute quantities.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

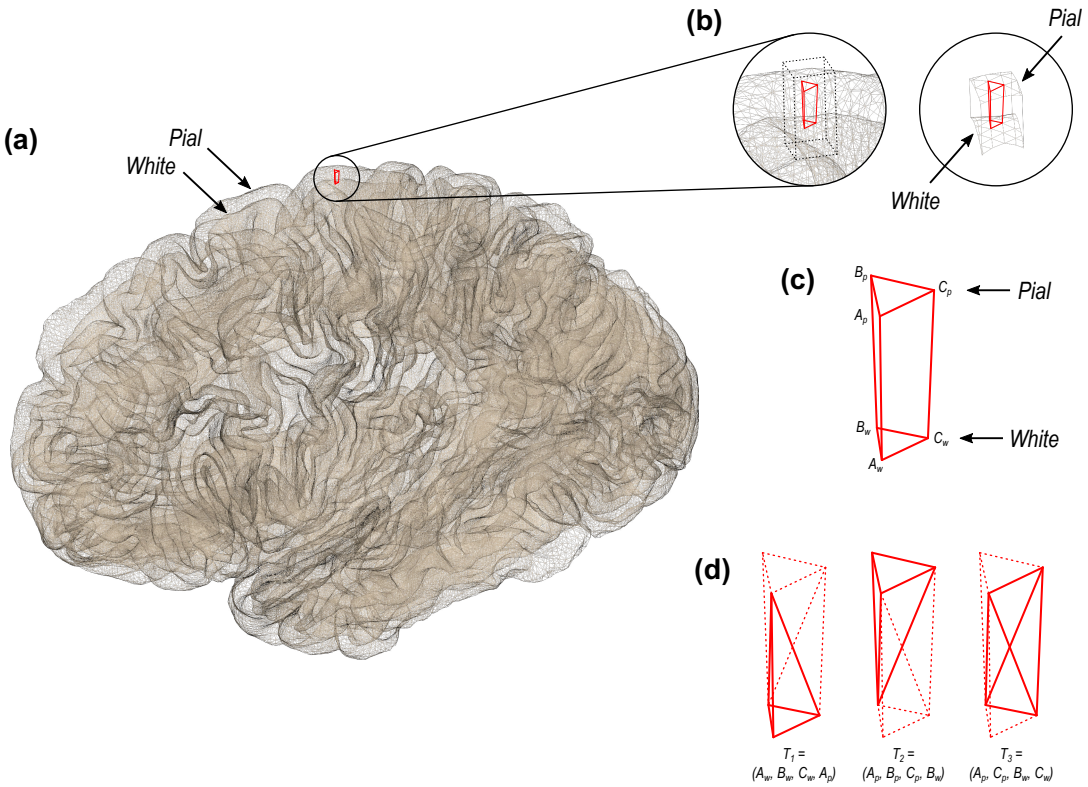


Figure 2: (a) In the surface representation, the cortex is limited internally by the white and externally by the pial surface. (b) and (c) These two surfaces have matching vertices that can be used to delineate an oblique truncated triangular pyramid. (d) The six vertices of this pyramid can be used to define three tetrahedra, the volumes of which are computed analytically.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Likewise, cerebral blood flow and volume obtained using methods based on magnetic resonance imaging (MRI), such as arterial spin labelling (ASL), as well as other forms of quantitative MRI, as those involving contrast enhancement (Parker and Padhani, 2003), quantitative magnetisation transfer (Levesque et al., 2010; Harrison et al., 2015), or quantitative assessment of myelination, are also areal quantities that require mass conservation when measured in absolute terms. The methods used for statistical analysis surface area can be applied for these areal quantities as well.

1.3. Non-parametric combination (NPC)

Instead of volume, a joint approach can be considered for the analysis of thickness and area. Classical multivariate tests such as MANCOVA, however, do not inform the direction of effects, and are based on assumptions that are known not to hold for surface area, such as normality. Rather, the permutation-based non-parametric combination (NPC) (Pesarin and Salmaso, 2010; Winkler et al., 2016b) provides a test for directional as well as two-tailed hypotheses, and is based on minimal assumptions, mainly that of exchangeability, that is, swapping one datum for another keeps the data just as likely.

The NPC consists of, in a first phase, testing separately hypotheses on each available metric using permutations that are performed in synchrony; these tests are termed *partial tests*. The resulting statistics for each and every permutation are recorded, allowing an estimate of the complete empirical cumulative distribution function (cdf) to be constructed for each one. In a second phase, the empirical p-values for each test are combined, for each permutation, into a *joint statistic*. As the joint statistic is produced from the previous permutations, an estimate of its empirical cdf function is immediately known, and so is its corresponding p-value.

As originally proposed and as described above, NPC is not practicable in brain imaging: as the statistics for all partial tests for all permutations need to be recorded, an enormous amount of space for data storage is necessary. However, even if storage space were not a problem, the discreteness of the p-values for the partial tests is problematic when correcting for multiple testing, because with thousands of vertices in a surface, ties occur frequently,

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

further causing ties among the combined statistics. If too many tests across an image share the same most extreme statistic, correction for the multiplicity, while still valid, is less powerful (Westfall and Young, 1993; Pantazis et al., 2005). The most obvious workaround — run an ever larger number of permutations to break the ties — may not be possible for small sample sizes, or when possible, requires correspondingly larger data storage. The solution to this problem is loosely based on the direct combination of the test statistics, by converting the statistics of the partial tests to values that behave as p-values using the asymptotic distribution of the statistics, and using these for the combination (Winkler et al., 2016b).

Combining functions. The null hypothesis of the NPC is that the null hypotheses for all partial tests are true, and the alternative is that any test is false, which is the same that a union-intersection test (UIT) (Roy, 1953). The rejection region depends on how the combined statistic is produced. Various combining functions can be considered, particularly those used in meta-analyses, such as Fisher's combination of p-values, i.e., $T = -2\sum_k \ln(p_k)$ (Fisher, 1932) and Stouffer's combination of z-statistics, i.e., $T = \sum_k \Phi^{-1}(1-p_k)/\sqrt{K}$ (Stouffer et al., 1949), where T is the test statistic for the joint test, p_k is the p-value of the k-th out of K partial tests, and Φ^{-1} is the probit function. These and most other combining functions, related statistics and their distributions were originally derived under the assumption of independence among the partial tests, which is not always valid, particularly under the tenable hypothesis of shared environmental effects affecting both area and thickness. Such lack of independence is not a problem for NPC: the synchronised permutations implicitly capture the dependencies among the tests that would cause a parametric combination to be invalid, even if using the same combining functions.

2. Method

The general workflow for surface-based morphometry consists of the generation of a surface-representation of the cortex and its subsequent homeomorphic transformation into a sphere. Vertices of this sphere are shifted

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

tangentially along its surface to allow alignment matching a particular feature of interest of a reference brain (i.e., an atlas), such as sulcal depth, myelin content, or functional markers. Once registration has been done, interpolation to a common grid is performed; it is at the resolution of this grid that analyses across subjects are performed. While the order of these processing stages remains generally fixed, the stage in which areal quantities are calculated or obtained varies according to the method: for the nearest neighbour, redistributive, and pycnophylactic methods, these are computed in the native space, using native geometry. With the retessellation method, area is computed in native space, with a new geometry produced after interpolation of the surface coordinates to the common grid. An overview of the whole process is in Figure 3.

We evaluate (1) if and where the four different interpolation methods (nearest neighbour, retessellation, redistributive and pycnophylactic) differ; (2) if and where these methods vary according to the resolution of the common grid used as target; (3) if and where the two ways of measuring volumes (the product method and the analytic method) differ from each other; (4) how the NPC results would relate to the separate analyses of thickness, area and volume.

2.1. Subjects

In the period 1986–88, 121 preterm newborns with very low birth weight (VLBW; $\leqslant$ 1500g) were admitted to the Neonatal Intensive Care Unit at the St. Olav University Hospital in Trondheim, Norway. At age 20, a total of 41 VLBW subjects consented to participate and had usable MRI data. The term-born controls were born at the same hospital in the same period. A random sample of women with parities 1 or 2 was selected for follow-up during pregnancy. At birth, 122 children with birth weight above the tenth percentile for gestational age from this sample were included as controls. At age 20, a total of 59 control subjects consented to participate and had usable MRI data. Further details can be found in Martinussen et al. (2005); Skranes et al. (2007).

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

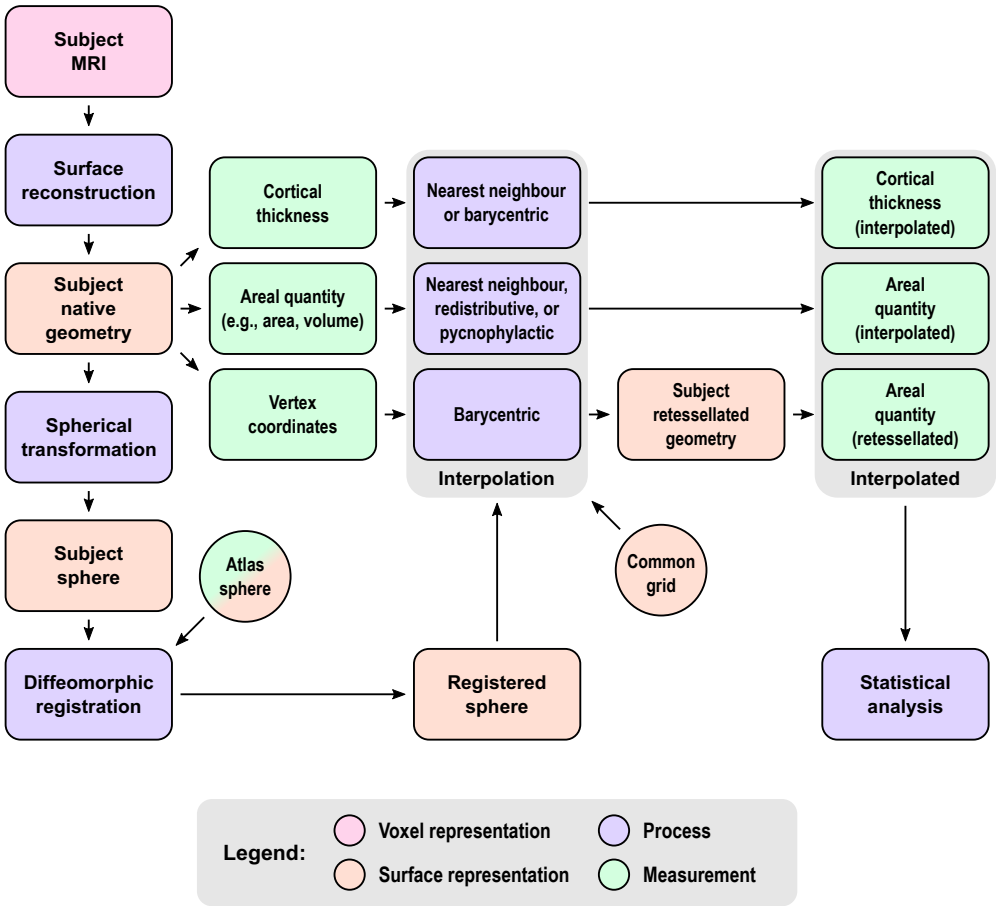


Figure 3: An overview of the steps for the analysis of surface area using different methods. The subject magnetic resonance images are used to reconstruct a pair of surfaces (pial and white) representing the cortex, which initially are in the subject space and individual geometry. From this pair of surfaces, cortical thickness can be measured. From the same surfaces, area and volume can be measured. Finally, the coordinates of the vertices can be stored for subsequent use. The subject native surfaces are homeomorphically transformed to a sphere, registered to a spherical atlas, and used for the interpolation, which for thickness can be either nearest neighbour or barycentric, for area can be nearest neighbour, redistributive or pycnophylactic, and for the vertex coordinates can be barycentric. In the latter, the interpolation of coordinates allows the construction of a new retessellated surface in subject space, from which area can alternatively be measured. The interpolated quantities are then ready to undergo statistical analyses. See references in the main text.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2.2. Data acquisition

MRI scanning was performed on a 1.5 T Siemens MAGNETOM Symphony scanner equipped with a quadrature head coil. In each scanning session, two sagittal T_1-weighted magnetization prepared rapid gradient echo (MPRAGE) scans/sequences were acquired (echo time = 3.45 ms, repetition time = 2730 ms, inversion time = 1000 ms, flip angle = 7°; field of view = 256 mm, voxel size = $1 \times 1 \times 1.33$ mm, acquisition matrix $256 \times 192 \times 128$).

2.3. Reconstruction of the cortical surface

We used the method implemented in the FreeSurfer software package (version 5.3.0; Dale et al., 1999; Fischl et al., 1999a): T_1-weighted images are first corrected for magnetic field inhomogeneities and then skull-stripped (Ségonne et al., 2004). Voxels belonging to the white matter (WM) are identified based on their locations, on their intensities, and on the intensities of the neighbouring voxels. A mass of connected WM voxels is produced for each hemisphere, using a six-neighbours connectivity scheme, and a mesh of triangular faces is tightly built around this mass, using two triangles for each externally facing voxel side. The mesh is smoothed taking into account the local intensity in the original images (Dale and Sereno, 1993), at a subvoxel resolution. Defects are corrected (Fischl et al., 2001; Ségonne et al., 2007) to ensure that the surface has the same topological properties of a sphere. A second iteration of smoothing is applied, resulting in a realistic representation of the interface between gray and white matter (the *white surface*). The external cortical surface (the *pial surface*), which corresponds to the pia mater, is produced by nudging outwards the white surface towards a point where the tissue contrast is maximal, between gray matter and CSF, maintaining constraints on its smoothness while preventing self-intersection. Cortical thickness is measured as the distance between the matching vertices of these two surfaces (Fischl and Dale, 2000).

2.4. Measurement of areal quantities

Areal quantities are measured in native space, i.e., before spherical transformation and registration. For the retessellation method, the measurement

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

is made in native space after the surface has been reconstructed to a particular resolution; for nearest neighbour, redistributive, and pycnophylactic, measurement uses native space, with the original, subject-specific mesh geometry.

Cortical area. For a triangular face ABC of the surface representation, with vertex coordinates $\mathbf{a} = [x_A\ y_A\ z_A]'$, $\mathbf{b} = [x_B\ y_B\ z_B]'$, and $\mathbf{c} = [x_C\ y_C\ z_C]'$, the area is $|\mathbf{u} \times \mathbf{v}|/2$, where $\mathbf{u} = \mathbf{a} - \mathbf{c}$, $\mathbf{v} = \mathbf{b} - \mathbf{c}$, $\times$ represents the cross product, and the bars $|\ |$ represent the vector norm. Although the area per face (i.e., the *facewise* area) can be used in subsequent steps, it remains the case that most software packages can only deal with values assigned to each vertex of the mesh (i.e., *vertexwise*). Conversion from facewise to vertexwise is achieved by assigning to each vertex one-third of the sum of the areas of all faces that have that vertex in common (Winkler et al., 2012).

Cortical volume. The conventional method for computing surface-based volume consists of computing the area at each vertex as above, then multiplying this value by the thickness at that vertex, in a procedure that leaves tissue under- or over-represented in gyri and sulci (Fig.1). Instead, volumes can be computed using the three vertices that define a face in the white surface and the three matching vertices in the pial surface, defining an *oblique truncated triangular pyramid*, which in turn is perfectly subdivided into three tetrahedra. The volumes of these are computed analytically, summed, and assigned to each face of the surface representation, viz.:

1. For a given face $A_wB_wC_w$ in the white surface, and its corresponding face $A_pB_pC_p$ in the pial surface, define an oblique truncated triangular pyramid.

2. Split this truncated pyramid into three tetrahedra, defined as:

$$\begin{aligned} T_1 &= (\ A_w,\ B_w,\ C_w,\ A_p\ ) \\ T_2 &= (\ A_p,\ B_p,\ C_p,\ B_w\ ) \\ T_3 &= (\ A_p,\ C_p,\ C_w,\ B_w\ ) \end{aligned}$$

   This division leaves no volume under- or over-represented.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

3. For each such tetrahedra, let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ represent its four vertices in terms of coordinates $[x\ y\ z]'$. Compute the volume as $|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|/6$, where $\mathbf{u} = \mathbf{a} - \mathbf{d}$, $\mathbf{v} = \mathbf{b} - \mathbf{d}$, $\mathbf{w} = \mathbf{c} - \mathbf{d}$, the symbol $\times$ represents the cross product, $\cdot$ represents the dot product, and the bars $|\ |$ represent the vector norm.

Computation can be accelerated by setting $\mathbf{d} = A_p$, the common vertex for the three tetrahedra, such that the vector subtractions can happen only once. Conversion from facewise volume to vertexwise is possible, and done in the same manner as for facewise area. The above method will be the default in the next FreeSurfer release.

2.5. Spherical transformation

The white surface is homeomorphically transformed to a sphere (Fischl et al., 1999b), thus keeping a one-to-one mapping between faces and vertices of the native geometry (white and pial) and the sphere. All these surfaces comprise triangular faces exclusively. Measurements of interest obtained from native geometry or in native space, such as area and thickness, are stored separately and are not affected by the transformation, nor by registration (next step; see diagram in *SI §1*).

2.6. Registration

Various strategies are available to put all surfaces in register and allow comparisons between, including the one used by FreeSurfer (Fischl et al., 1999b), Spherical Demons (SD) (Yeo et al., 2010), Multimodal Surface Matching (MSM) (Robinson et al., 2014), among others. From these, methods that are diffeomorphic (i.e., smooth and invertible) should be favoured. Methods that are not diffeomorphic by construction, but that in practice produce invertible and smooth warps can, in principle, be used for registration for areal analyses. For these analyses, the FreeSurfer method was used; a side comparison with SD is in the Supplementary Information).

2.7. Interpolation methods

Statistical comparisons require meshes with a common resolution where each point (vertex, face) represents homologous regions across individuals.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A mesh that can act as a common grid is a geodesic sphere constructed by iterative subdivision of the faces of a regular (Platonic) icosahedron. A geodesic sphere has many advantages as the target for interpolation: ease of computation, edges of roughly similar sizes and, if the resolution is fine enough, edge lengths that are much smaller than the diameter of the sphere (Kenner, 1976). We compared four different interpolation methods each at three different mesh resolutions: IC3 (lowest resolution, with 642 vertices and 1280 faces), IC5 (intermediate resolution, with 10242 vertices and 20480 faces), and IC7 (163842 vertices and 327680 faces).

Nearest neighbour interpolation. The well known nearest neighbour interpolation does not guarantee preservation of areal quantities, although modifications can be introduced to render it approximately mass conservative: for each vertex in the target, the closest vertex is found in the source sphere, and the area from the source vertex is assigned to the target vertex; if a given source vertex maps to multiple target vertices, its area is divided between them so as to preserve the total area. If there are any source vertices that have not been represented in the target, for each one of these, the closest target vertex is located, and the corresponding area from the source surface is incremented to any area already stored on it. This method ensures that the total area after mapping into the group surface remains unchanged. This process is a surface equivalent of Jacobian correction[2] used in volume-based methods in that it accounts for stretches and shrinkages while preserving the overall amount of areal quantities. The nearest neighbour is currently the default method in FreeSurfer.

Retessellation of the native geometry. This method appeared in Saad et al. (2004). It consists of generating a new mesh by interpolating not the area assigned to vertices, but the coordinates of the corresponding vertices in the native geometry. The set of three coordinates is used, together with

[2]Not to be confused with the computation of the Jacobian itself, that is computed, for the i-th vertex, as $J_i = \frac{A_i^S}{A_i^w} \frac{\sum_i A_i^w}{\sum_i A_i^S}$, where A_i^S is the area of the vertex in the source (registered) sphere, A_i^w is the area of the same vertex in the white surface (native space and native geometry), and the sums are over the entire surface, i.e., all vertices.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the connectivity scheme between vertices from the common grid, to construct a new mesh that has overall similar shape as the original brain, but with the geometry of the common grid. The areas for each face (or vertex) can be computed from this new mesh, and therefore be used for statistical comparisons among subjects. Equivalently, the coordinates of each vertex can be treated as a single vector, and the barycentric interpolation can be performed in a single step as:

$$\begin{bmatrix} x_P \\ y_P \\ z_P \end{bmatrix} = \begin{bmatrix} x_A & x_B & x_C \\ y_A & y_B & y_C \\ z_A & z_B & z_C \end{bmatrix} \begin{bmatrix} \delta_A \\ \delta_B \\ \delta_C \end{bmatrix}$$

where x, y, z represent the coordinates of the triangular face ABC and of the interpolated point P, both in native geometry, and δ are the barycentric coordinates of P with respect to the same face after the spherical transformation. From the four methods considered in this paper, this is the only one that does not directly interpolate either area or areal quantities, but the mesh in native space.

Redistribution of areas. This method works by splitting the areal quantity present in each vertex from the source sphere using the proportion given by the barycentric coordinates of that vertex in relation to the face in the target sphere (common grid) on which it lies, redistributing these quantities to the three vertices that constitute that face in the target. If some quantity was already present in these target vertex (e.g., from other source vertices lying on the same target face), that is incremented. The method is represented by:

$$Q_i^T = \sum_{f=1}^{F} \sum_{v=1}^{V_f} Q_{vf}^S \delta_{ivf}$$

where Q_{vf}^S is the areal quantity in the source vertex v, $v \in \{1, \ldots, V_f\}$ lying on the target face f, $f \in \{1, \ldots, F\}$, F being the number of faces that meet at the target vertex i, and δ_{ivf} is the barycentric coordinate of v, lying on face f, and in relation to the target vertex i. This method has similarities with the conventional barycentric interpolation (as used for the

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

interpolation of coordinates in the retesselation method). The key difference is that in the barycentric interpolation, it is the barycentric coordinates of the target vertex in relation to their containing source face that are used to weight the quantities, in a process that therefore is not mass conservative. Here it is the barycentric coordinates of the source vertex in relation to their containing target face that are used; the quantities are split proportionately, and redistributed across target vertices.

Pycnophylactic interpolation. The ideal interpolation method should conserve the areal quantities globally, regionally and locally. In other words, the method has to be *pycnophylactic*. This is accomplished in this method by assigning, to each face in the target sphere, the areal quantity of all overlapping faces from the source sphere, weighted by the fraction of overlap between them (Markoff and Shapiro, 1973; Winkler et al., 2012). It operates on the faces directly, not on vertices. The area (or any other areal quantity) is transferred from source to target surface via weighting by the overlapping area between any pairs of faces. The interpolated areal quantity, Q_i^T, of a face i in the target surface, that overlaps with F faces from the source surface, is given by:

$$Q_i^T = \sum_{f=1}^{F} \frac{A_f^O}{A_f^S} Q_f^S$$

where A_f^S is the area of the f-th overlapping face from the source sphere, which contains a quantity Q_f^S of some areal measurement (such as the surface area measured in the native space), and A_f^O is the overlapping area with the face i.

Correction for unequal face sizes and smoothing. Regardless of the interpolation method, larger faces in the common grid inherit larger amounts of areal quantities. If the analysis will compare regions that are topographically distinct, or if the data are to be smoothed, a correction for different face sizes is needed (Winkler et al., 2012). For facewise data, it consists of weighting the the areal quantity at each face or vertex, after interpolation, by a constant that depends on the respective area in the common grid. Smoothing was considered at two levels for the comparison of areal

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

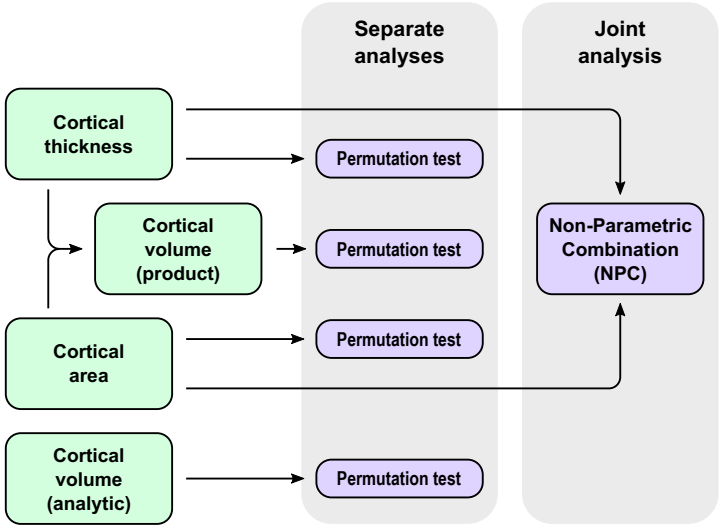


Figure 4: Overview of the separate and joint analyses of thickness, area and volume.

interpolation and volume methods: no smoothing, and smoothing with a Gaussian kernel with full width at half maximum (FWHM) of 10 mm, small so as to preserve the effect of different resolutions being investigated. For the comparison between VLBW and controls, 30 mm, as in Skranes et al. (2013).

2.8. Statistical analysis

The statistical analysis was performed using PALM – Permutation Analysis of Linear Models (available at fsl.fmrib.ox.ac.uk) (Winkler et al., 2014, 2016b). The number of permutations was set at 1000, followed by approximation of the tail of the distribution by a generalised Pareto distribution (GPD) (Winkler et al., 2016a), and familywise error rate correction (FWER) was applied considering both hemispheres and both test directions for the null hypothesis of no difference between both subject groups. Analyses were performed separately for cortical thickness, area, and volume (both methods), and also using NPC for the joint analysis of thickness and area (Figure 4.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

3. Results

3.1. Preservation of areal quantities

At the highest resolution of the common grid used as the interpolation target (IC7), all methods preserve generally well the global amount of surface area, and therefore, of other areal quantities. At lower resolutions, massive amounts of area are lost with the retessellation method: about 40% on average for IC3 and 9% for IC5, although only 1% for IC7. Areal losses, when existing, tend to be uniformly distributed across the cortex (Fig.5, upper panels), with no trends affecting particular regions and, except for retessellation, can be substantially alleviated by smoothing (*SI §2*).

3.2. Differences between interpolation methods

While there are no spatial trends in terms of areal gains or losses, the inexactness of some interpolation methods introduces noise that substantially reduces their correlation when assessed between subjects (Fig.5, lower panels). The only exception is between the retessellation and the pycnophylactic methods, which have near perfect correlation even without any smoothing. Smoothing increases the correlation between all methods to near unity throughout the cortex (*SI §2a*). At the subject level, the spatial correlation between the nearest neighbour and the pycnophylactic methods is only about 0.60, although approaching unity when the subjects are averaged (*SI §2b*). Smoothing leads to a dramatic improvement on agreement, causing nearest neighbour to be nearly indistinguishable from the pycnophylactic method. The redistributive method performed in a similar manner, although with a higher correlation without smoothing, about 0.75 (*SI §2b*).

3.3. Cortical volume measurements

At the local scale, differences between the product and the analytic methods to estimate volume is as high as 20% in some regions (*SI §3*), an amount that could not be alleviated by smoothing or by changes in resolution. As predicted by Fig.1, differences were larger in the crowns of gyri and depths of sulci, in either case with the reverse polarity (Fig.6, upper panels). The vertexwise correlation between the methods across subjects, however, was in general very high, approaching unity throughout the whole cortex, with or

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

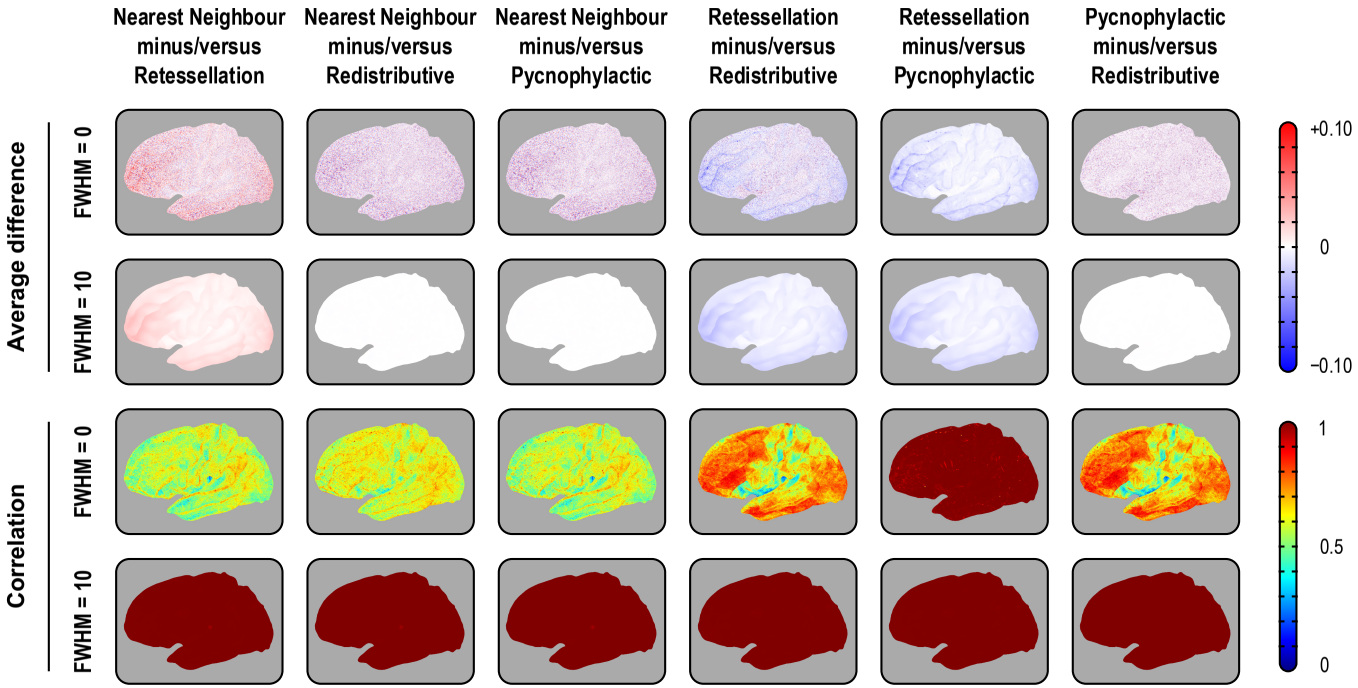


Figure 5: Pairwise average differences (in mm^2) and correlations between the four interpolation methods, using the IC7 as target, with or without smoothing with a Gaussian kernel of FWHM = 10 mm, projected to the average white surface. Although the four methods differ, with some leading to substantial, undesirable losses and gains in surface area, and the introduction of noise manifested by lower correlations, the average variation was zero for nearest neighbour, redistributive and pycnophylactic. The retessellation method led to substantial losses of area that could not be recovered or compensated by blurring. Although this method showed excellent correlation with pycnophylactic, quantitative results after interpolation are biased downwards. For the medial views, for the right hemisphere, for IC3 and IC5, and for projections to the pial and inflated surfaces, consult the Supplemental Material.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

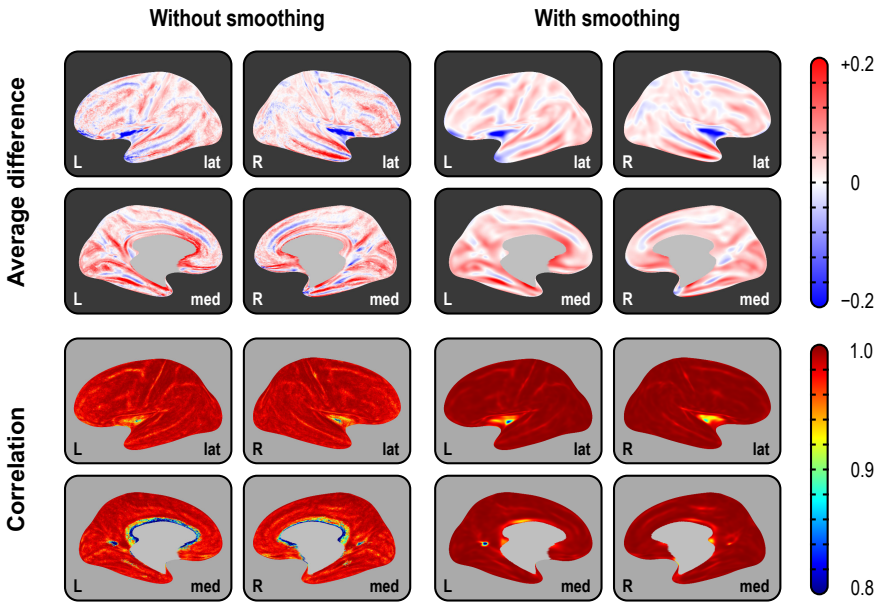


Figure 6: Average difference (in mm^3) between the two methods of assessing volume and their correlation (across subjects), using the highest resolution (IC7) as the interpolation target, projected to the average inflated surface. As predicated from Fig. 1, differences are larger in the crowns of gyri and in the depths of sulci, with gains/losses in volume in these locations following opposite patterns. Although the correlations tend to be generally high, and increase with smoothing, they are lower in regions of higher inter-individual morphological variability, such as at the anterior end of the cuneus, and in the insular cortex. For IC3 and IC5, and for projections to the white and pial surfaces, consult the Supplemental Material.

without smoothing, and at different resolutions. In regions of higher sulcal variability, however, the correlations were not as high, sometimes as low as 0.80, such as in the insular cortex and at the confluence of parieto-occipital and calcarine sulci, between the lingual and the isthmus of the cingulate gyrus (Fig.6, lower panels). At least in the case of the insula, this effect may be partly attributed to a misplacement of the white surface in the region lateral to the claustrum (Glasser et al., 2016).

3.4. Variability of thickness, area and volume

Regardless of the methods, variability of area is higher than for thickness, and even higher for volume: the average coefficient of variation across subjects ($100 \cdot \sigma/\mu$) was, respectively, 9.9%, 3.2% and 10.5%, after adjusting for group, age, and sex, with the parietal region (bilateral) being the most

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

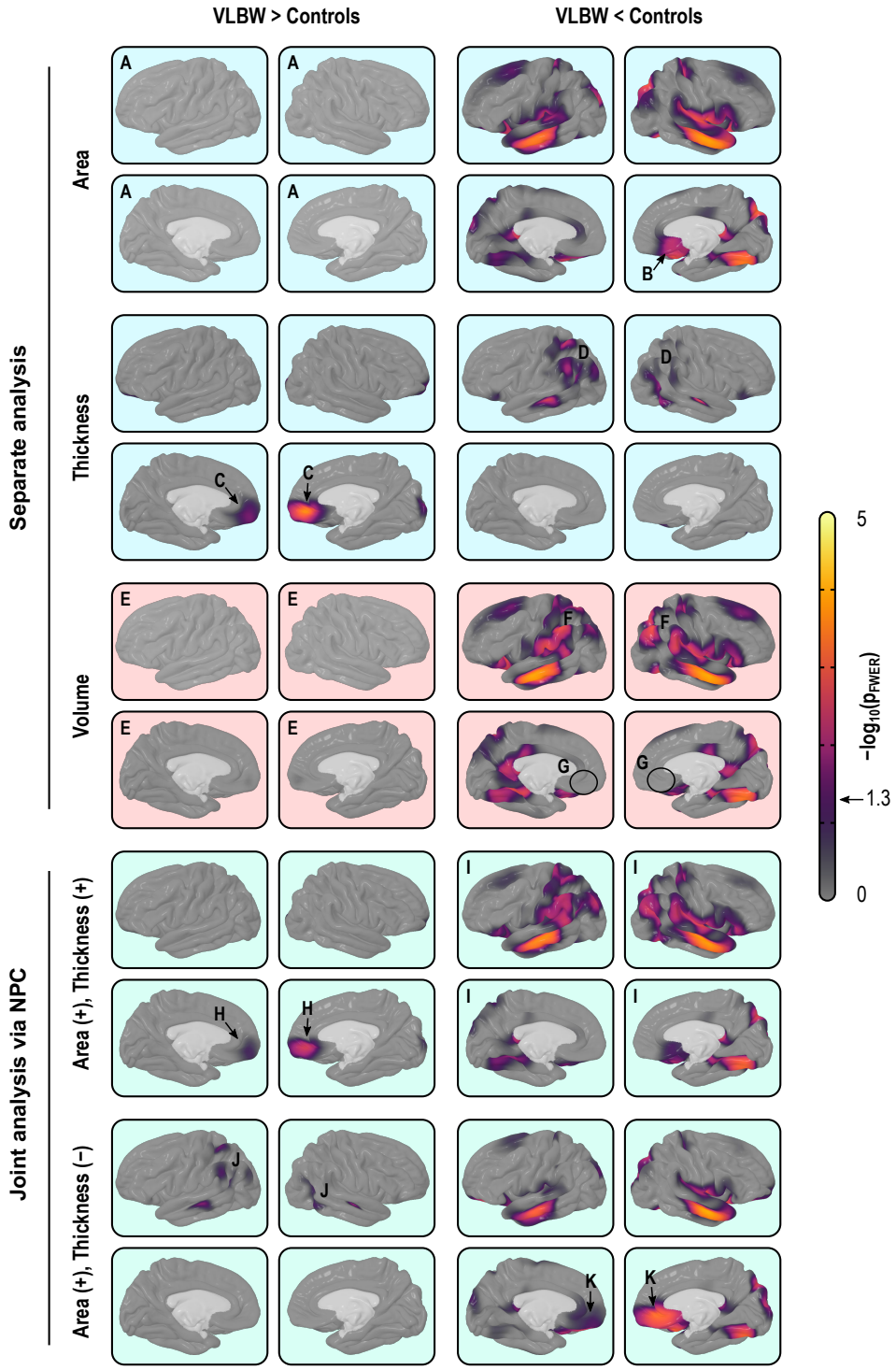
variable for all measurements (for spatial maps, see *SI §4*).

3.5. Differences between VLBW *and controls*

Analysing cortical thickness and area separately, the comparisons between the VLBW subjects and the controls suggest a distinct pattern of significant differences ($p \leqslant 0.05$, FWER-corrected). Surface area maps show a significant bilateral reduction in the middle temporal gyrus, the superior banks of the lateral sulcus, and the occipito-temporal lateral (fusiform) gyrus, as well as a diffuse bilateral pattern of areal losses affecting the superior frontal gyrus, posterior parietal cortex and, in the right hemisphere, the subgenual area of the cingulate cortex. Cortical thickness maps suggest a diffuse bilateral thinning in the parietal lobes, left middle temporal gyrus, right superior temporal sulcus, and bilateral thickening of the medial orbito-frontal cortex of the VLBW subjects compared to controls (Fig.7, upper panels, light blue background). Maps of cortical volume differences largely mimic the surface area results, although with a few differences: diffuse signs of volume reduction in the parietal lobes, ascribable to cortical thinning and, contrary to the analysis of area and thickness, no effects found in the medial-orbitofrontal or in the subgenual region of the cingulate gyrus (Fig.7, middle panels, light red background).

Figure 7: Separate (*light blue background*) and joint (*green*) analysis of cortical area and thickness, as well as volume (*red*), using the IC7 resolution and smoothing with FWHM = 30mm. Analysis of area indicates no reductions in the control group anywhere in the cortex (A), and among other regions, the subgenual region of the cingulate cortex (B). Analysis of thickness indicates that VLBW subjects have thicker cortex in the medial orbito-frontal cortex (C) and diffuse bilateral thinning mostly in the parietal and middle temporal regions (D). Analysis of volume alone broadly mimics analysis of area, with no evidence for increased volume in VLBW subjects (E), although some maps there seems to be a partial superimposition, with signs of bilateral decreased volume in the parietal lobe (F), but differently than for the analysis of area, no signs for reductions in the subgenual cortex (G). Jointly analysing area and thickness gives equal weight to both measurements, and allows directional effects to be inferred. Differently than in the case for volume, it is possible to know that there is an increase in the amount of cortical tissue in VLBW subjects in the medial orbito-frontal cortex (H) when compared to controls, and a bilateral decrease throughout most of the parietal cortex, more strongly in the middle temporal and fusiform gyri, in both hemispheres (I). Moreover, the joint analysis allows search for effects that can negate each other, such as in this case weaker effects in the parietal region (J), that partially overlap in space with those shown in (I). Finally, strong effects in the middle orbito-frontal, that were missed with simple volumes (G) become clearly visible (K).

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

3.6. Joint analysis via NPC

Non-parametric combination of thickness and area provides information about patterns of group differences not visible in cortical volume analyses, or that appear split or not visible in separate maps of area and thickness (Fig.7, lower panels, light green background). In the present data, the joint analysis suggests a decrease in the amount of tissue in VLBW subjects in the medial orbito-frontal cortex, and a bilateral decrease throughout most of the parietal cortex, as well as in the middle temporal and fusiform gyri. Furthermore, it suggests weaker bilateral increase in the amount of tissue in the parietal region, that alternates in space with that of tissue loss. Finally, NPC shows simultaneous bilateral decrease in surface area and increase in thickness in the medial orbito-frontal gyrus, none of which was observed using simple volume measurements (for additional maps, see *SI §5*).

4. Discussion

4.1. Interpolation of areal quantities

All the interpolation approaches examined can be applied to other areal quantities, including those measured in volume-based representations of the brain. In these cases, quantities are projected from each voxel to the nearest vertex or face of the native geometry, leaving no voxel with valid measurements unrepresented. Once the data has been projected (transferred), the volume representation is no longer necessary, and nearest neighbour, redistributive and pycnophylactic methods can be considered. For the retessellation method, the retessellated surface is produced first, and it is to this surface that the quantities from the volumetric representation are transferred, either voxelwise or facewise.

However, and irrespective of the data, the different interpolation methods assessed do not perform similarly in all settings. The nearest neighbour and redistributive methods require smoothing in order to become comparable to, and interchangeable with, the pycnophylactic method. Since data is usually smoothed in neuroimaging studies in order to improve matching of homologies between subjects and to improve the signal-to-noise ratio, this is not in practice a limitation. Retesselation, particularly at lower resolutions,

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

leads to substantial areal losses that cannot be recovered even after smoothing. Moreover, the vertices of the retessellated surfaces are not guaranteed to lie at the tissue boundaries they aim to represent, introducing uncertainties on obtained measurements. Regarding speed, although the various implementations run on linear time, $\mathcal{O}(n)$, the pycnophylactic method has to perform a larger number of computations that may not pay off when compared with nearest neighbour, provided that smoothing is used.

4.2. Areal expansion and absolute area

Few studies of cortical surface area have offered insight into the procedures adopted. Sometimes the methods were described as areal expansion/contraction, as opposed to measuring surface area itself. Furthermore, different definitions of areal expansion/contraction have been used, e.g., relative to the contra-lateral hemisphere (Lyttelton et al., 2009), to some earlier point in time (Hill et al., 2010), to a control group (Palaniyappan et al., 2011), or in relation to a standard brain, possibly the default brain (average or atlas) used in the respective software package (Joyner et al., 2009; Rimol et al., 2010a, 2012; Chen et al., 2011, 2012; Vuoksimaa et al., 2016); other studies considered linear distances as proxies for expansion/contraction (Sun et al., 2009a,b). Some of the studies that used a default brain as reference did use nearest neighbour interpolation followed by smoothing (Joyner et al., 2009; Rimol et al., 2010a, 2012), which as we have shown, assesses cortical area itself; nevertheless, the measurements were described in terms of areal expansion/contraction. These multiple definitions make the interpretation and comparison between studies challenging.

Whether regional proportions within the neocortex would be relatively constant across primates (Schoenemann et al., 2005; Barton and Venditti, 2013; Gabi et al., 2016) is an example question that can be answered by the use of areal quantities, conditional on the definition of borders. Measurements of areal expansion/contraction in relation to a given reference can also be obtained once interpolation has been performed using the methods presented here. It suffices to divide the area per face (or per vertex) by the area of the corresponding face (or vertex) in the brain used as reference, which can be an atlas, the contralateral hemisphere, the same brain at an

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

earlier point in time, or a brain from a different species. Surface area and areal expansion/contraction are related to each other by a factor that varies spatially.

4.3. Volumes improved, yet problematic

The large absolute difference between the product and analytic methods for measuring cortical volume indicates that if interest lies in the actual values (for instance, for predictive models), the proposed, analytic method for assessing volume is to be preferred. The high correlation across subjects, however, suggests that, for group comparisons and similar analyses, both methods lead to generally similar results, except in a few regions of higher morphological inter-individual variability. However, even in these cases, cortical volume is a poor choice of trait of interest, since it is largely insensitive to changes in cortical thickness, and it is easily replaceable by superior alternatives. While volume encapsulates information on both area and thickness, research has suggested that the proportion in which the variability of these two measurements coalesces varies spatially across the cortical mantle (Winkler et al., 2010), and moreover, that most of the variability of cortical volume, including that measured using VBM, can be explained by the variability of surface area alone (Voets et al., 2008; Lenroot et al., 2009; Winkler et al., 2010; Rimol et al., 2012), rendering volume a largely redundant metric. The continuous cortical maps in in Fig.7 provide evidence that the results for gray matter volume, even without using the simple product of thickness by area, mostly mirror the results for cortical surface area. However, the two are not interchangeable, as cortical thickness has some influence, such that in the absence of volume effects, the possibility that changes in area have been compensated by changes in thickness, or vice versa, to no net effect, cannot be excluded.

4.4. Joint analyses via NPC

Such problems with cortical volume can be eschewed through the use of a joint statistical analysis of area and thickness. The NPC methodology gives equal (or otherwise predefined) weights for thickness and area, which

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

therefore do not mix in unknown and variable proportions across the cortical mantle. Various combining functions can be considered, and the well known Fisher method of combination of p-values is suggested as a simple and computationally efficient choice (Fisher, 1932). By using two distinct metrics in a single test, power is increased (Pesarin and Salmaso, 2010; Winkler et al., 2016b), allowing detection of effects that otherwise may remain unseen when analysing volume, or when thickness and area are used separately. NPC can be particularly useful for the investigation of processes affecting cortical area and thickness simultaneously, and can effectively replace volume as the measurement of interest in these cases, with various beneficial effects, and essentially none of the shortcomings. It constitutes a general method that can be applied to any number of partial tests, each relating to hypotheses on data that may be of different nature, obtained using different measurement units, and related to each other arbitrarily.

Moreover, NPC allows testing directional hypotheses (by reversing the signs of partial tests), hypotheses with concordant directions (taking the extremum of both after multiple testing correction), and two-tailed hypotheses (with two-tailed partial tests). Power increases consistently with the introduction of more partial tests when there is a true effect, while strictly controlling the error rate. This is in stark contrast with classical multivariate tests based on regression, such as MANOVA or MANCOVA, that do not provide information on directionality of the effects, and lose power as the number of partial tests increase past a certain maximum point (Winkler et al., 2016b).

A joint test using NPC has similarities with, yet it is distinct from, the test known as *conjunction* or *intersection-union test* (IUT) (Nichols et al., 2005). The NPC tests a joint null hypothesis that all partial tests have no effect; if the null is rejected in any partial test at a suitable level, the joint null is rejected. The conjunction tests a null hypothesis that at least one partial test has no effect; the alternative is that all partial tests have an effect. While a conjunction seeks an effect across all tests, NPC seeks an effect in any, or in an aggregate of the partial tests. Usage of NPC is not constrained to the replacement of cortical volume, and the method can be considered for analyses involving other cortical indices, including myelination (Glasser and Van Essen, 2011; Sereno et al., 2013) and folding and gyrification metrics

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(Mangin et al., 2004; Schaer et al., 2008; Toro et al., 2008) among others.

4.5. Permutation inference

Permutation tests provide exact inference based on minimal assumptions, while allowing multiple testing correction with strong control over the error rate. Even though permutation tests still have certain requirements, such that the data are exchangeable, certain types of structured dependency can be accommodated by means of restricted permutation strategies. Being based on permutations in each of the partial tests, NPC does not preclude the analysis of thickness and area (or of other partial tests) separately, and through the synchronised shuffling, correction for multiplicity of tests while taking into account their non-independence is trivial. This includes correction for multiple tests that may be used using various combinations of positive and negative directions for the partial tests. Permutation tests do not depend on distributional assumptions, which favours the analysis of surface area, which at the local level shows positive skewness, and is better characterised as log-normal (Winkler et al., 2012).

4.6. Area and thickness of VLBW subjects

The reduced cortical surface area observed in the VLBW subjects compared to controls replicates previous findings from the same cohort at 20 years of age (Skranes et al., 2013), and is consistent with findings from a younger cohort of VLBW subjects (Sølsnes et al., 2015), and teenagers born with extremely low birth weight ($\leqslant$ 1000g) (Grunewaldt et al., 2014). The combined evidence from these studies suggests that surface area reductions in the preterm brain are present from early childhood and remain reasonably constant from childhood until adulthood (Rimol et al., 2016). Proposed mechanisms for gray matter injury in preterm birth include hypoxia-ischemia and inflammation arising from intrauterine infections or from postnatal sepsis (Volpe, 2009, 2011), which may adversely affect critical phases of brain maturation before and after birth and cause diffuse white matter damage, including hypomyelination and primary or secondary gray matter dysmaturation (Hagberg et al., 2015). Cortical area reductions may not be explained by primary white matter damage alone, especially since area reductions are

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

also observed in younger cohorts of preterms with less perinatal morbidity and less pathology in white matter microstructure evaluated with diffusion tensor imaging (Eikenes et al., 2011; Rimol et al., 2016). Reduced neuropil is a possible explanation for cortical thinning in the lateral parietal and temporal cortex in VLBW subjects, but the thickening of the medial orbito-frontal cortex must be due to different mechanisms (Marín-Padilla, 1997; Bjuland et al., 2013; Grunewaldt et al., 2014). The combination of thickening and reduced area in medial orbito-frontal cortex has been observed in multiple cohorts, and more generally, these changes in both thickness and area could be related both to prenatal factors, such as foetal growth restriction, or to postnatal exposure to extra-uterine environmental stressors (Sølsnes et al., 2015; Rimol et al., 2016). Regardless of the underlying pathological aspects, the morphological indices appear to be robust markers of perinatal brain injury and maldevelopment (Raznahan et al., 2011; Skranes et al., 2013; Rimol et al., 2016), and the effects observed replicate earlier findings.

Supplementary information

The large number of scenarios evaluated, that involved two different registration and four different interpolation methods, three grid resolutions, two different smoothing levels, four different indices of cortical morphology, plus NPC, resulted in more than 16 thousand maps and Bland-Altman plots (Bland and Altman, 1986). These have been organised in a set of browsable pages accessible at http://bit.ly/2cHJFQC. These constitute the Supplementary Information, and their inspection is encouraged. The SI also includes high resolution and complementary views of all figures shown in the main text.

Acknowledgements

The authors thank Donald J. Hagler Jr. (University of California, San Diego, CA, USA), and David Van Essen and Matthew F. Glasser (Washington University Medical School, St. Louis, MO, USA) for the helpful discussions at different stages of this work. A.M.W. is supported by the National Research Council of Brazil (CNPq; 211534/2013-7). T.E.N. is supported by the Wellcome Trust (100309/Z/12/Z). M.R.S. received funding from the NIBIB/NIH K25-EB013649. Data processing was largely performed on the Abel Cluster, owned by the University of Oslo and the Norwegian Metacenter for High Performance Computing (NOTUR), and operated by the Department for Research Computing at USIT, University of Oslo.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

Ashburner, J., 2009. Computational anatomy with the SPM software. Magn Reson Imaging 27 (8), 1163–74.

Ashburner, J., Friston, K. J., 2000. Voxel-based morphometry - the methods. Neuroimage 11, 805–21.

Barton, R. A., Venditti, C., 2013. Human frontal lobes are not relatively large. Proc Natl Acad Sci USA 110 (22), 9001–6.

Bjuland, K. J., Løhaugen, G. C. C., Martinussen, M., Skranes, J., 2013. Cortical thickness and cognition in very-low-birth-weight late teenagers. Early Hum Dev 89 (6), 371–380.

Bland, J. M., Altman, D. G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327 (8476), 307–10.

Ceccarelli, A., Rocca, M. A., Pagani, E., Colombo, B., Martinelli, V., Comi, G., Filippi, M., 2008. A voxel-based morphometry study of grey matter loss in MS patients with different clinical phenotypes. Neuroimage 42 (1), 315–322.

Chen, C.-H., Gutierrez, E. D., Thompson, W., Panizzon, M. S., Jernigan, T. L., Eyler, L. T., Fennema-Notestine, C., Jak, A. J., Neale, M. C., Franz, C. E., Lyons, M. J., Grant, M. D., Fischl, B., Seidman, L. J., Tsuang, M. T., Kremen, W. S., Dale, A. M., 2012. Hierarchical genetic organization of human cortical surface area. Science 335 (6076), 1634–6.

Chen, C.-H., Panizzon, M. S., Eyler, L. T., Jernigan, T. L., Thompson, W., Fennema-Notestine, C., Jak, A. J., Neale, M. C., Franz, C. E., Hamza, S., Lyons, M. J., Grant, M. D., Fischl, B., Seidman, L. J., Tsuang, M. T., Kremen, W. S., Dale, A. M., 2011. Genetic influences on cortical regionalization in the human brain. Neuron 72 (4), 537–544.

Dale, A. M., Fischl, B., Sereno, M. I., 1999. Cortical surface-based analysis I: Segmentation and surface reconstruction. Neuroimage 9 (2), 179–94.

Dale, A. M., Sereno, M. I., 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction. J Cogn Neurosci 5 (2), 162–176.

Douaud, G., Smith, S., Jenkinson, M., Behrens, T., Johansen-Berg, H., Vickers, J., James, S., Voets, N., Watkins, K., Matthews, P. M., James, A., 2007. Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. Brain 130 (9), 2375–2386.

Eikenes, L., Løhaugen, G. C., Brubakk, A. M., Skranes, J., Håberg, A. K., 2011. Young adults born preterm with very low birth weight demonstrate widespread white matter alterations on brain DTI. Neuroimage 54 (3), 1774–1785.

Fischl, B., Dale, A. M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc Natl Acad Sci USA 97 (20), 11050–5.

Fischl, B., Liu, A., Dale, A. M., 2001. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. IEEE Trans Med Imaging 20 (1), 70–80.

Fischl, B., Sereno, M. I., Dale, A. M., 1999a. Cortical surface-based analysis II: Inflation, flattening, and a surface-based coordinate system. Neuroimage 9 (2), 195–207.

Fischl, B., Sereno, M. I., Tootell, R. B., Dale, A. M., 1999b. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum Brain Mapp 8 (4), 272–84.

Fisher, R. A., 1932. Statistical Methods for Research Workers. Oliver and Boyd, Edinburgh.

Fjell, A. M., Grydeland, H., Krogsrud, S. K., Amlien, I., Rohani, D. A., Ferschmann, L., Storsve, A. B., Tamnes, C. K., Sala-Llonch, R., Due-Tønnessen, P., Bjørnerud, A., Sølsnes, A. E., Håberg, A. K., Skranes, J., Bartsch, H., Chen, C.-H., Thompson, W. K., Panizzon, M. S., Kremen, W. S., Dale, A. M., Walhovd, K. B., dec 2015. Development and aging of cortical thickness correspond to genetic organization patterns. Proc Natl Acad Sci USA 112 (50), 15462–15467.

Gabi, M., Neves, K., Masseron, C., Ribeiro, P. F. M., Ventura-Antunes, L., Torres, L., Mota, B., Kaas, J. H., Herculano-Houzel, S., 2016. No relative expansion of the number of prefrontal neurons in primate and human evolution. Proc Natl Acad Sci USA 113 (34), 9617–22.

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Geschwind, D. H., Rakic, P., 2013. Cortical evolution: judge the brain by its cover. Neuron 80 (3), 633–47.

Gitelman, D. R., Ashburner, J., Friston, K. J., Tyler, L. K., Price, C. J., 2001. Voxel-based morphometry of herpes simplex encephalitis. Neuroimage 13 (4), 623–31.

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., Van Essen, D. C., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536 (7615), 171–8.

Glasser, M. F., Van Essen, D. C., 2011. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. J Neurosci 31 (32), 11597–616.

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., Frackowiak, R. S., 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage 14 (1 Pt 1), 21–36.

Grunewaldt, K. H., Fjørtoft, T., Bjuland, K. J., Brubakk, A.-M., Eikenes, L., Håberg, A. K., Løhaugen, G. C. C., Skranes, J., 2014. Follow-up at age 10 years in ELBW children - functional outcome, brain morphology and results from motor assessments in infancy. Early Hum Dev 90 (10), 571–8.

Hagberg, H., Mallard, C., Ferriero, D. M., Vannucci, S. J., Levison, S. W., Vexler, Z. S., Gressens, P., 2015. The role of inflammation in perinatal brain injury. Nat Rev Neurol 11 (4), 192–208.

Harrison, N. A., Cooper, E., Dowell, N. G., Keramida, G., Voon, V., Critchley, H. D., Cercignani, M., 2015. Quantitative magnetization transfer imaging as a biomarker for effects of systemic inflammation on the brain. Biol Psychiatry 78 (1), 49–57.

Hill, J., Inder, T., Neil, J., Dierker, D., Harwell, J., van Essen, D., 2010. Similar patterns of cortical expansion during human development and evolution. Proc Natl Acad Sci USA 107 (29), 13135–40.

Hogstrom, L. J., Westlye, L. T., Walhovd, K. B., Fjell, A. M., 2013. The structure of the cerebral cortex across adult life: age-related patterns of surface area, thickness, and gyrification. Cereb Cortex 23 (11), 2521–30.

Joyner, A. H., Roddey, J. C., Bloss, C. S., Bakken, T. E., Rimol, L. M., Melle, I., Agartz, I., Djurovic, S., Topol, E. J., Schork, N. J., Andreassen, O. A., Dale, A. M., 2009. A common MECP2 haplotype associates with reduced cortical surface area in humans in two independent populations. Proc Natl Acad Sci USA 106 (36), 15483–8.

Kenner, H., 1976. Geodesic math and how to use it. University of California Press, Los Angeles, CA, USA.

Küper, M., Rabe, K., Esser, S., Gizewski, E. R., Husstedt, I. W., Maschke, M., Obermann, M., 2011. Structural gray and white matter changes in patients with HIV. J Neurol 258 (6), 1066–1075.

Leahy, R. M., Qi, J., 2000. Statistical approaches in quantitative positron emission tomography. Stat Comput 10 (2), 147–165.

Lenroot, R. K., Schmitt, J. E., Ordaz, S. J., Wallace, G. L., Neale, M. C., Lerch, J. P., Kendler, K. S., Evans, A. C., Giedd, J. N., 2009. Differences in genetic and environmental influences on the human cerebral cortex associated with development during childhood and adolescence. Hum Brain Mapp 30 (1), 163–74.

Levesque, I. R., Giacomini, P. S., Narayanan, S., Ribeiro, L. T., Sled, J. G., Arnold, D. L., Pike, G. B., 2010. Quantitative magnetization transfer and myelin water imaging of the evolution of acute multiple sclerosis lesions. Magn Reson Med 63 (3), 633–640.

Lyttelton, O. C., Karama, S., Ad-Dab'bagh, Y., Zatorre, R. J., Carbonell, F., Worsley, K., Evans, A. C., 2009. Positional and surface area asymmetry of the human cerebral cortex. Neuroimage 46 (4), 895–903.

Mangin, J.-F., Rivière, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Scifo, P., Ochiai, T., Brunelle, F., Régis, J., 2004. A framework to study the cortical folding patterns. NeuroImage 23 Suppl 1, S129–38.

Marín-Padilla, M., 1997. Developmental neuropathology and impact of perinatal brain damage. II: white matter lesions of the neocortex. J Neuropathol Exp Neurol 56 (3), 219–35.

Markoff, J., Shapiro, G., 1973. The linkage of data describing overlapping geographical units. Hist Methods Newsl 7 (1), 34–46.

Marsland, A. L., Gianaros, P. J., Abramowitch, S. M., Manuck, S. B., Hariri, A. R., 2008. Interleukin-6 covaries inversely with hippocampal grey matter volume in middle-aged

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

adults. Biol Psychiatry 64 (6), 484–490.

Martinussen, M., Fischl, B., Larsson, H. B., Skranes, J., Kulseng, S., Vangberg, T. R., Vik, T., Brubakk, A. M., Haraldseth, O., Dale, A. M., 2005. Cerebral cortex thickness in 15-year-old adolescents with low birth weight measured by an automated MRI-based method. Brain 128 (11), 2588–2596.

Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.-B., 2005. Valid conjunction inference with the minimum statistic. NeuroImage 25 (3), 653–60.

Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., Kan, E., Kuperman, J. M., Akshoomoff, N., Amaral, D. G., Bloss, C. S., Libiger, O., Schork, N. J., Murray, S. S., Casey, B. J., Chang, L., Ernst, T. M., Frazier, J. A., Gruen, J. R., Kennedy, D. N., Van Zijl, P., Mostofsky, S., Kaufmann, W. E., Kenet, T., Dale, A. M., Jernigan, T. L., Sowell, E. R., 2015. Family income, parental education and brain structure in children and adolescents. Nat Neurosci 18 (5), 773–778.

O'Leary, D. D. M., Chou, S.-J., Sahara, S., 2007. Area patterning of the mammalian cortex. Neuron 56 (2), 252–69.

Palaniyappan, L., Mallikarjun, P., Joseph, V., White, T. P., Liddle, P. F., 2011. Regional contraction of brain surface area involves three large-scale networks in schizophrenia. Schizophr Res 129 (2-3), 163–8.

Panizzon, M. S., Fennema-Notestine, C., Eyler, L. T., Jernigan, T. L., Prom-Wormley, E., Neale, M., Jacobson, K., Lyons, M. J., Grant, M. D., Franz, C. E., Xian, H., Tsuang, M., Fischl, B., Seidman, L., Dale, A., Kremen, W. S., 2009. Distinct genetic influences on cortical surface area and cortical thickness. Cereb Cortex 19 (11), 2728–35.

Pantazis, D., Nichols, T. E., Baillet, S., Leahy, R. M., 2005. A comparison of random field theory and permutation methods for the statistical analysis of MEG data. Neuroimage 25 (2), 383–94.

Parker, G. J. M., Padhani, A. R., 2003. T_1-w DCE-MRI: T_1-weighted Dynamic Contrast-Enhanced MRI. In: Tofts, P. (Ed.), Quantitative MRI of the Brain: Measuring Changes Caused by Disease. John Wiley & Sons, pp. 341–364.

Pesarin, F., Salmaso, L., 2010. Permutation Tests for Complex Data: Theory, Applications and Software. John Wiley and Sons, West Sussex, England, UK.

Rakic, P., 1988. Specification of cerebral cortical areas. Science 241 (4862), 170–6.

Rakic, P., sep 1995. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. Trends Neurosci 18 (9), 383–8.

Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G. L., Greenstein, D., Clasen, L., Gogtay, N., Giedd, J. N., 2011. How does your cortex grow? J Neurosci 31 (19), 7174–7.

Rimol, L. M., Agartz, I., Djurovic, S., Brown, A. A., Roddey, J. C., Kähler, A. K., Mattingsdal, M., Athanasiu, L., Joyner, A. H., Schork, N. J., Halgren, E., Sundet, K., Melle, I., Dale, A. M., Andreassen, O. A., Alzheimer's Disease Neuroimaging Initiative, 2010a. Sex-dependent association of common variants of microcephaly genes with brain structure. Proc Natl Acad Sci USA 107 (1), 384–8.

Rimol, L. M., Bjuland, K. J., Løhaugen, G. C. C., Martinussen, M., Evensen, K. A. I., Indredavik, M. S., Brubakk, A.-M., Eikenes, L., Håberg, A. K., Skranes, J., 2016. Cortical trajectories during adolescence in preterm born teenagers with very low birthweight. Cortex 75, 120–31.

Rimol, L. M., Nesvå g, R., Hagler, D. J., Bergmann, O., Fennema-Notestine, C., Hartberg, C. B., Haukvik, U. K., Lange, E., Pung, C. J., Server, A., Melle, I., Andreassen, O. A., Agartz, I., Dale, A. M., 2012. Cortical volume, surface area, and thickness in schizophrenia and bipolar disorder. Biol Psychiatry 71 (6), 552–60.

Rimol, L. M., Panizzon, M. S., Fennema-Notestine, C., Eyler, L. T., Fischl, B., Franz, C. E., Hagler, D. J., Lyons, M. J., Neale, M. C., Pacheco, J., Perry, M. E., Schmitt, J. E., Grant, M. D., Seidman, L. J., Thermenos, H. W., Tsuang, M. T., Eisen, S. A., Kremen, W. S., Dale, A. M., 2010b. Cortical thickness is influenced by regionally specific genetic factors. Biol Psychiatry 67 (5), 493–9.

Robinson, E. C., Jbabdi, S., Glasser, M. F., Andersson, J., Burgess, G. C., Harms, M. P., Smith, S. M., Van Essen, D. C., Jenkinson, M., 2014. MSM: A new flexible framework for multimodal surface matching. Neuroimage 100, 414–26.

Roy, S. N., Jun. 1953. On a heuristic method of test construction and its use in multivariate analysis. Ann Math Stat 24 (2), 220–238.

Saad, Z. S., Reynolds, R. C., Argall, B., Japee, S., Cox, R. W., 2004. SUMA: An inter-

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

face for surface-based intra- and inter-subject analysis with AFNI. IEEE International Symposium on Biomedical Imaging, 1510–1513.

Schaer, M., Cuadra, M. B., Tamarit, L., Lazeyras, F., Eliez, S., Thiran, J.-P., 2008. A surface-based approach to quantify local cortical gyrification. IEEE Trans Med Imaging 27 (2), 161–70.

Schmitt, J. E., Lenroot, R. K., Wallace, G. L., Ordaz, S., Taylor, K. N., Kabani, N. J., Greenstein, D., Lerch, J. P., Kendler, K. S., Neale, M. C., Giedd, J. N., 2008. Identification of genetically mediated cortical networks: a multivariate study of pediatric twins and siblings. Cereb Cortex 18 (8), 1737–47.

Schnack, H. G., Van Haren, N. E. M., Brouwer, R. M., Evans, A., Durston, S., Boomsma, D. I., Kahn, R. S., Hulshoff Pol, H. E., 2015. Changes in thickness and surface area of the human cortex and their relationship with intelligence. Cereb Cortex 25 (6), 1608–1617.

Schoenemann, P. T., Sheehan, M. J., Glotzer, L. D., 2005. Prefrontal white matter volume is disproportionately larger in humans than in other primates. Nat Neurosci 8 (2), 242–52.

Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., Fischl, B., 2004. A hybrid approach to the skull stripping problem in MRI. Neuroimage 22 (3), 1060–75.

Ségonne, F., Pacheco, J., Fischl, B., 2007. Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. IEEE Trans Med Imaging 26 (4), 518–29.

Sereno, M. I., Lutti, A., Weiskopf, N., Dick, F., 2013. Mapping the human cortical surface by combining quantitative T1 with retinotopy. Cereb Cortex 23 (9), 2261–8.

Skranes, J., Løhaugen, G. C. C., Martinussen, M., Håberg, A., Brubakk, A. M., Dale, A. M., 2013. Cortical surface area and IQ in very-low-birth-weight (VLBW) young adults. Cortex 49 (8), 2264–2271.

Skranes, J., Vangberg, T. R., Kulseng, S., Indredavik, M. S., Evensen, K. A. I., Martinussen, M., Dale, A. M., Haraldseth, O., Brubakk, A. M., 2007. Clinical findings and white matter abnormalities seen on diffusion tensor imaging in adolescents with very low birth weight. Brain 130 (3), 654–666.

Sølsnes, A. E., Grunewaldt, K. H., Bjuland, K. J., Stavnes, E. M., Bastholm, I. a., Aanes, S., Østgård, H. F., Håberg, A., Løhaugen, G. C. C., Skranes, J., Rimol, L. M., 2015. Cortical morphometry and IQ in VLBW children without cerebral palsy born in 2003-2007. Neuroimage. Clinical 8, 193–201.

Stouffer, S. A., Suchman, E. A., DeVinney, L. C., Star, S. A., Jr., R. M. W., 1949. The American Soldier: Adjustment During Army Life (Volume 1). Princeton University Press, Princeton, New Jersey, USA.

Sun, D., Phillips, L., Velakoulis, D., Yung, A., McGorry, P. D., Wood, S. J., van Erp, T. G. M., Thompson, P. M., Toga, A. W., Cannon, T. D., Pantelis, C., 2009a. Progressive brain structural changes mapped as psychosis develops in 'at risk' individuals. Schizophr Res 108 (1-3), 85–92.

Sun, D., Stuart, G. W., Jenkinson, M., Wood, S. J., McGorry, P. D., Velakoulis, D., van Erp, T. G. M., Thompson, P. M., Toga, a. W., Smith, D. J., Cannon, T. D., Pantelis, C., 2009b. Brain surface contraction mapped in first-episode schizophrenia: a longitudinal magnetic resonance imaging study. Mol Psychiatry 14 (10), 976–86.

Toro, R., Perron, M., Pike, B., Richer, L., Veillette, S., Pausova, Z., Paus, T., 2008. Brain size and folding of the human cerebral cortex. Cereb Cortex 18 (10), 2352–7.

van den Hoff, J., 2005. Principles of quantitative positron emission tomography. Amino Acids 29 (4), 341–53.

van Essen, D. C., 2005. A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. Neuroimage 28 (3), 635–62.

Voets, N. L., Hough, M. G., Douaud, G., Matthews, P. M., James, A., Winmill, L., Webster, P., Smith, S., 2008. Evidence for abnormalities of cortical development in adolescent-onset schizophrenia. Neuroimage 43 (4), 665–75.

Volpe, J. J., 2009. Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. Lancet Neurol 8 (1), 110–124.

Volpe, J. J., 2011. Systemic inflammation, oligodendroglial maturation, and the encephalopathy of prematurity. Ann Neurol 70 (4), 525–529.

Vuoksimaa, E., Panizzon, M. S., Chen, C.-H., Fiecas, M., Eyler, L. T., Fennema-Notestine, C., Hagler, D. J., Franz, C. E., Jak, A. J., Lyons, M. J., Neale, M. C., Rinker, D. A., Thompson, W. K., Tsuang, M. T., Dale, A. M., Kremen, W. S., 2016. Is bigger al-

bioRxiv preprint doi: https://doi.org/10.1101/074666; this version posted September 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ways better? the importance of cortical configuration with respect to cognitive ability. Neuroimage 129, 356–366.

Westfall, P. H., Young, S. S., 1993. Resampling-Based Multiple Testing: Examples And Methods for p-Value Adjustment. John Wiley and Sons, New York.

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., Duggirala, R., Glahn, D. C., 2010. Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. Neuroimage 15 (3), 1135–46.

Winkler, A. M., Ridgway, G. R., Douaud, G., Nichols, T. E., Smith, S. M., 2016a. Faster permutation inference in brain imaging. NeuroImage 141, 502–516.

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., Nichols, T. E., 2014. Permutation inference for the general linear model. Neuroimage 92, 381–97.

Winkler, A. M., Sabuncu, M. R., Yeo, B. T. T., Fischl, B., Greve, D. N., Kochunov, P., Nichols, T. E., Blangero, J., Glahn, D. C., 2012. Measuring and comparing brain cortical surface area and other areal quantities. Neuroimage 61 (4), 1428–43.

Winkler, A. M., Webster, M. A., Brooks, J. C., Tracey, I., Smith, S. M., Nichols, T. E., 2016b. Non-Parametric Combination and related permutation tests for neuroimaging. Hum Brain Mapp 37 (4), 1486–511.

Yeo, B. T., Sabuncu, M. R., Vercauteren, T., Ayache, N., Fischl, B., Golland, P., 2010. Spherical demons: fast diffeomorphic landmark-free surface registration. IEEE Trans Med Imaging 29 (3), 650–68.

Zhang, H., Sachdev, P. S., Wen, W., Crawford, J. D., Brodaty, H., Baune, B. T., Kochan, N. A., Slavin, M. J., Reppermund, S., Kang, K., Trollor, J. N., 2015. The relationship between inflammatory markers and voxel-based gray matter volumes in nondemented older adults. Neurobiol Aging 37, 138–146.

Zhang, Z., Wang, Y., Shen, Z., Yang, Z., Li, L., Chen, D., Yan, G., Cheng, X., Shen, Y., Tang, X., Hu, W., Wu, R., 2016. The neurochemical and microstructural changes in the brain of systemic lupus erythematosus patients: A multimodal MRI study. Scientific Reports 6, 19026.