NeuroImage 63 (2012) 365–380

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Benefits of multi-modal fusion analysis on a large-scale dataset: Life-span patterns of inter-subject variability in cortical morphometry and white matter microstructure

Adrian R. Groves [a,*], Stephen M. Smith [a], Anders M. Fjell [b], Christian K. Tamnes [b], Kristine B. Walhovd [b], Gwenaëlle Douaud [a], Mark W. Woolrich [a,c], Lars T. Westlye [b,**]

[a] FMRIB (Oxford Centre for Functional Magnetic Resonance Imaging of the Brain), Nuffield Department of Clinical Neurosciences, University of Oxford, OX3 9DU, UK
[b] Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Norway, PO Box 1094 Blindern, 0317 Oslo, Norway
[c] Oxford Centre for Human Brain Activity, Department of Psychiatry, University of Oxford, UK

ARTICLE INFO

Article history:
Accepted 20 June 2012
Available online 29 June 2012

Keywords:
Independent component analysis
Data fusion
Diffusion tensor imaging
Voxel-based morphometry
Surface-based morphometry
Bayesian modeling
Aging

ABSTRACT

Neuroimaging studies have become increasingly multimodal in recent years, with researchers typically acquiring several different types of MRI data and processing them along separate pipelines that provide a set of complementary windows into each subject's brain. However, few attempts have been made to integrate the various modalities in the same analysis. Linked ICA is a robust data fusion model that takes multi-modal data and characterizes inter-subject variability in terms of a set of multi-modal components. This paper examines the types of components found when running Linked ICA on a large magnetic resonance imaging (MRI) morphometric and diffusion tensor imaging (DTI) data set comprising 484 healthy subjects ranging from 8 to 85 years of age. We find several strong global features related to age, sex, and intracranial volume; in particular, one component predicts age to a high accuracy ($r = 0.95$). Most of the remaining components describe spatially localized modes of variability in white or gray matter, with many components including both tissue types. The multimodal components tend to be located in anatomically-related brain areas, suggesting a morphological and possibly functional relationship. The local components show relationships between surface-based cortical thickness and arealization, voxel-based morphometry (VBM), and between three different DTI measures. Further, we report components related to artifacts (e.g. scanner software upgrades) which would be expected in a dataset of this size. Most of the 100 extracted components showed interpretable spatial patterns and were found to be reliable using split-half validation. This work provides novel information about normal inter-subject variability in brain structure, and demonstrates the potential of Linked ICA as a feature-extracting data fusion approach across modalities. This exploratory approach automatically generates models to explain structure in the data, and may prove especially powerful for large-scale studies, where the population variability can be explored in increased detail.

Introduction

Genetic, environmental and lifestyle interactions are constantly molding our brains in different ways, effectively shaping the between-subject differences in behavioral and brain measures. In line with the notion of multi-level and interacting modulating factors on brain structure, functionally relevant differences in clinical and normal populations are often hypothesized to be *intrinsically multimodal*, including both gray and white matter measures (Zatorre et al., 2012). Thus, magnetic resonance (MR) neuroimaging has become increasingly multimodal in recent years, with many studies including multiple pulse sequences generating many different imaging contrasts from each subject. Functional, morphometric, and diffusion MRI data are typically analyzed through separate pipelines, and by jointly interpreting the separate results together one can gain a more comprehensive understanding of the brain. However, this approach is limited because the modalities are only combined at the point of interpretation. There should be benefits to fusing the data earlier, in particular after preprocessing but before statistical analysis, in order to model multi-modal covariances across subjects and characterize the between-subject variability in a data-driven way.

Exploratory methods that describe the variability across modalities are useful for a number of reasons. Since the various modalities are structural indices of the same subjects' brains, a high degree of common variance is assumed across modalities, and efficient modeling of this shared variance may increase sensitivity to behavioral, environmental, age-related or genetic variability. Seeley et al. (2009) showed that distinct patterns of neurodegeneration observed using voxel-based morphometry (VBM) (Good et al., 2001) in a variety of diseases can

* Corresponding author. Fax: +44 1865 222717.
** Corresponding author. Fax: +47 22845001.
E-mail addresses: adriang@fmrib.ox.ac.uk, adrian-neuroimage@groves.ca (A.R. Groves), l.t.westlye@psykologi.uio.no (L.T. Westlye).

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2012.06.038

also be reconstructed as patterns of VBM *variability* in a healthy control population. This suggests that the structured patterns found in healthy controls may also be relevant for decoding disease state. Multivariate data-driven methods may also be able to find subtle structured artifacts and automatically model these, effectively regressing confounds out of the data, further increasing sensitivity to biological variance. Finally, the components summarize the whole data set, reducing the dimensionality from many thousands of voxels (and their associated covariances) to a manageable number of individually interpretable, independent components.

Importantly, the forces driving the individual structural brain differences may modulate both specific and general processes, effectively influencing various levels or dimensions, e.g. through specific or general influences on cortical gray or white brain matter, respectively. Here, we attempt to study the patterns of inter-subject variability across imaging modalities using an exploratory tool for characterizing the statistical and, indirectly, the biological relationships between structural brain imaging modalities, and for pinpointing the remaining artifactual or morphological differences in these modalities. To this end we examine data from 484 subjects covering a wide age range (8–85 years). T1-weighted structural and diffusion MRI data from these subjects were analyzed using existing optimized modality-specific pipelines to obtain spatial maps of cortical morphometric measures including surface based measures of cortical thickness and cortical arealization (Dale et al., 1999, 2000; Fischl and Dale, 2000; Fischl et al., 1999a), voxel-based morphometry (VBM) based estimation of local gray matter volume/density (Douaud et al., 2007), as well as three white matter DTI indices: fractional anisotropy (FA), mean diffusivity (MD) and the mode of the diffusion tensor (MO) (Ennis and Kindlmann, 2006).

These measures were fused using a Linked ICA model (Groves et al., 2011) to obtain independent components of multimodal variability. Linked ICA automatically balances the information content of different modalities, finding subject loadings that produce statistically independent and non-Gaussian spatial maps across the modalities. This enables the modalities with the strongest information relating to each underlying feature to drive the subject-course of the component. To aid interpretability and draw meaningful conclusions, we correlate each component's subject-course with age, sex and intracranial volume (ICV). Other useful information is also obtained, such as the contribution of each modality to determining the component.

Sample and methods

Sample

The sample comprising 484 right-handed healthy volunteers covering the age range from 8 to 85 years old was drawn from the first wave of two longitudinal research projects at the Center for the Study of Human Cognition at the University of Oslo: *Neurocognitive Development* (Østby et al., 2009; Tamnes et al., 2010) and *Cognition and Plasticity through the Life-Span* (Fjell et al., 2008; Westlye et al., 2009a). The study was approved by the Regional Ethical Committee of Southern Norway (REK-Sør). The participants were recruited through newspaper advertisements and selected from among students and employees of the University of Oslo. We obtained written informed consent from all participants ≥12 years and from parents for participants <18 years of age. Oral informed consent was given by participants <12 years of age.

All subjects were native Norwegian speakers and screened using a standardized health interview prior to inclusion in the study. Subjects with a history of self- or parent-reported neurological or psychiatric conditions including clinically significant stroke, serious head injury, untreated hypertension, diabetes, and use of psychoactive drugs within the last two years were excluded. Further, participants reporting worries concerning their cognitive status, including memory function, were excluded. In addition, all included individuals' MR scans were examined by a neuroradiologist and deemed free of significant anomalies, including brain tumors and significant vascular insults (stroke). All subjects above 20 years of age scored <16 on Beck Depression Inventory (Beck and Steer, 1987) and subjects above 40 years of age ≥26 on Mini Mental State Examination (Folstein et al., 1975). General cognitive abilities were assessed by Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999). Mean IQ for the sample was 112.9 (SD = 10.2, range = 82–145).

Data acquisition and preparation

Imaging was performed using a 12-channel head coil on a 1.5-T Siemens Avanto scanner (Siemens Medical Solutions, Erlangen, Germany) at Oslo University Hospital, Rikshospitalet, with no hardware upgrades and only minor software upgrades performed during the course of the acquisition period (2006–2010).

The pulse sequence used for T1-weighted structural imaging was magnetization prepared rapid gradient echo (MP-RAGE), with the following parameters: TR/TE/time to inversion (TI)/flip angle (FA) = 2400 ms/3.61 ms/1000 ms/8°, matrix 192 × 192, field of view = 240, voxel size = 1.25 × 1.25 × 1.20 mm, and 160 sagittal slices. Scanning time was 7 min 42 s. Two runs were averaged during data processing to increase signal-to-noise-ratio (SNR).

For diffusion weighted imaging a single-shot twice-refocused spin-echo echo planar imaging pulse sequence with the following parameters was used: repetition time (TR)/echo time (TE) = 8200 ms/82 ms, b-value = 700 s/mm^2, voxel size = 2.0 × 2.0 × 2.0 mm, and 64 axial slices. The sequence was repeated in 2 successive runs with 10 images acquired with a b-value of 0, and 30 diffusion weighted images collected per run. Acquisition time was 11 min 21 s. This sequence is optimized to minimize eddy current-induced distortions (Reese et al., 2003). The two acquisitions were averaged during data processing to increase SNR. The protocol also included a 176 slices sagittal 3D T2-weighted turbo spin-echo sequence (TR/TE = 3390/388 ms), and a 25 slices coronal T2-FLAIR sequence (TR/TE = 7000–9000/109 ms) used for clinical assessment.

Optimized modality-specific preprocessing pipelines were used to produce standard-space subject images for each modality. Note that the alignment across subjects *within* each modality is essential, but there is no need for alignment *across* modalities; indeed, different resolutions and masks can be used, and voxel- and surface-based modalities are combined freely. Fig. 1 gives an overview of the pipelines used, to extract six maps of interest from the two scan types.

Preparation of DTI data using TBSS

Processing of DTI data was performed using the FMRIB Software Library, FSL (Smith et al., 2004; Woolrich et al., 2009). Each volume was affine registered to the first b0 volume using FMRIB's linear image registration tool FLIRT (Jenkinson et al., 2002) to correct for motion between images and eddy currents. After removal of non-brain tissue (Smith, 2002), FA, eigenvector, and eigenvalue maps were computed by linearly fitting a diffusion tensor to the data.

In addition to FA and MD, we include tensor mode (MO) because it is mathematically (though not biologically) orthogonal to FA and MD and allows separating planar from linear diffusion tensors (Ennis and Kindlmann, 2006). MO is particularly sensitive to subtle differences in regions with crossing fibers (Douaud et al., 2011). MD was defined as the mean of the eigenvalues.

FA volumes were skeletonized and transformed into common space (Smith et al., 2006, 2007). All volumes were warped to the FMRIB58_FA template using local deformation procedures performed by FMRIB's nonlinear image registration tool (FNIRT) (Andersson et al., 2007a,b). Excellent native-to-standard warping across individuals in a partly overlapping life-span sample was previously demonstrated (Westlye et al., 2010b). A mean FA volume of all individuals was

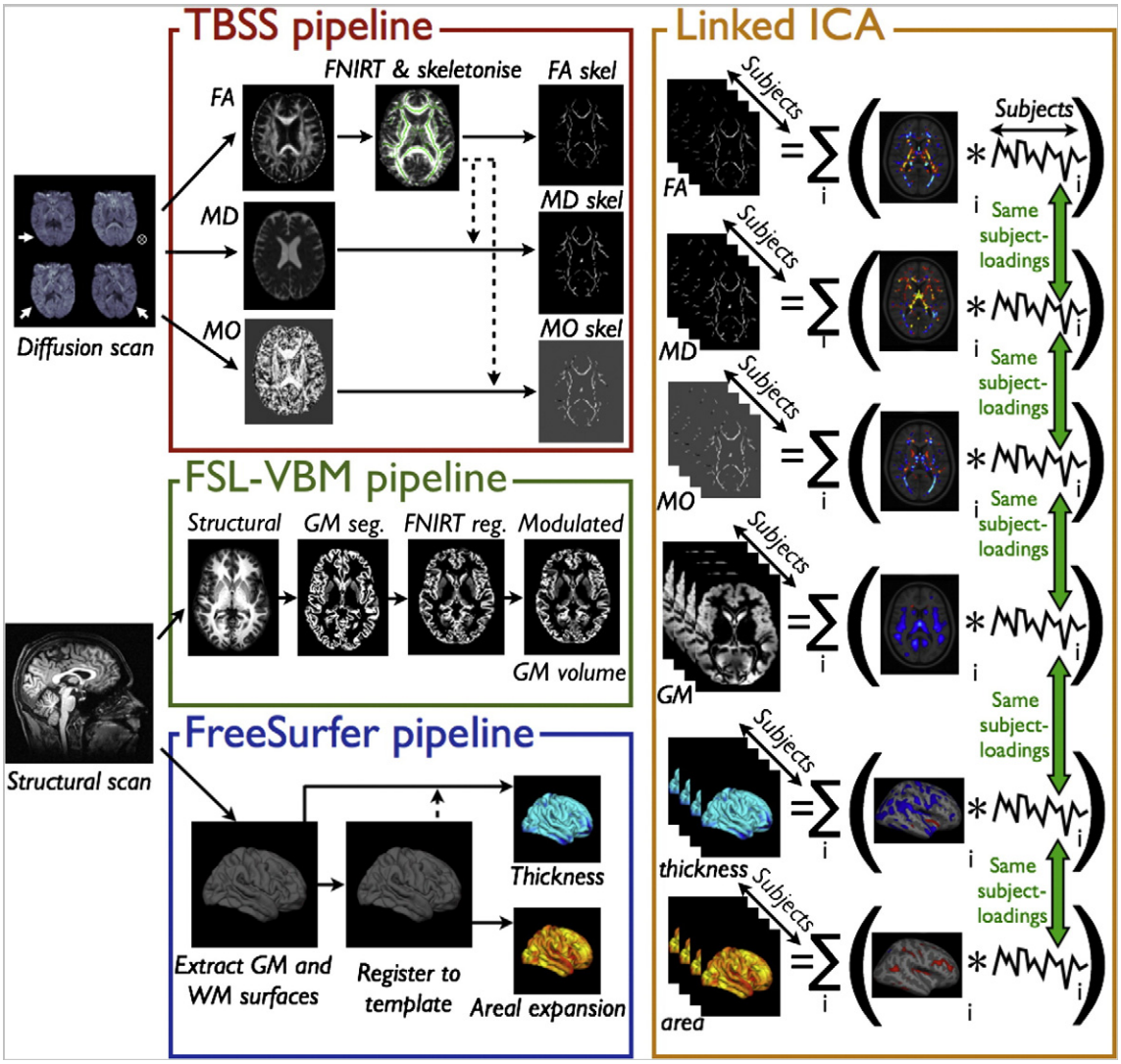

Fig. 1. A summary of the preprocessing pipelines, linked ICA, and post-processing steps used in this paper. The inputs are processed using standard TBSS, VBM, and FreeSurfer pipelines, then after Linked ICA analysis the resulting component spatial maps are rendered while the component subject-courses are correlated with various subject variables.

thinned to create a mean FA skeleton representing the centers of all tracts. We thresholded and binarized the mean skeleton at FA exceeding 0.2 to minimize partial voluming at the boundaries between tissue classes, yielding a skeleton of 127,562 voxels. Individual FA values were warped onto this mean skeleton by searching perpendicularly from the skeleton for maximum FA, further minimizing partial voluming. The same warping was applied to the MD and MO data, yielding skeletons sampled from voxels with FA exceeding 0.2.

Preparation of structural data using surface-based and volumetric approaches

Vertex-wise cortical thickness and "arealization" measures across the brain surface were estimated using FreeSurfer (http://surfer.nmr.mgh.harvard.edu/) by means of an automated surface reconstruction scheme (Dale et al., 1999; Fischl and Dale, 2000; Fischl et al., 1999a, 1999b, 2001, 2002, 2004).

Cortical thickness measurements were obtained by reconstructing representations of the gray/white boundary and the pial surface (Dale and Sereno, 1993; Dale et al., 1999) and then by calculating the distance between the surfaces at each vertex across the cortical mantle. Surface area was estimated by registering each subject's reconstructed surfaces to a common template (using folding information to drive the within-surface warping), and the relative amount of expansion or compression at each vertex was used as a proxy for regional arealization. This area measure is not normalized for head size. Surface maps were resampled, mapped to a common coordinate system using a non-rigid high-dimensional spherical averaging method to align cortical folding patterns (Fischl et al., 1999a, 1999b, 2008), and smoothed with a Gaussian kernel with a full width of half maximum of 10 mm.

T1-weighted data was also processed using FSL-VBM (Douaud et al., 2007), a VBM-based analysis (Ashburner and Friston, 2000; Good et al., 2001). The registered gray matter partial volume maps were modulated by dividing by the Jacobian of the warp field to correct for local expansion or contraction. The modulated segmented images were then smoothed with an isotropic Gaussian kernel with a sigma of 4 mm (FWHM = 9.4 mm).

Downsampling

For computational reasons, we reduced the number of data points in each modality. Since the FSL-VBM preprocessing already included a significant amount of smoothing (FWHM = 9.4 mm), downsampling from 2 mm to 4 mm isotropic causes minimal loss of information (with the trilinear interpolation increasing the effective smoothing to FWHM = 9.8 mm). However, it does make the images less clearly interpretable, so after running Linked ICA on the downsampled data

we then re-fit the final decomposition using the full-resolution data. The images shown in this paper are in the original, high-resolution space; the low-resolution versions are qualitatively identical.

We reduced the TBSS skeleton resolution to 2 mm isotropic, with the data projected from the 1 mm skeleton (127,350 voxels) onto a 2 mm skeleton (11,860 voxels) obtained by downsampling and renormalizing the 1 mm skeleton; this reduced the data dimensionality while ensuring that the same tracts were included in both resolutions of data. This also demonstrates that existing optimized pipelines can be used, without having to re-run analyses at a lower resolution just for linked ICA. FreeSurfer data was sampled from subject space directly to the high-resolution fsaverage surface (163,842 cortical vertices per hemisphere) and then to the low-resolution fsaverage5 template (10,242 vertices); both were then smoothed with a surface FWHM of 10 mm.

Linked ICA

Linked ICA is a data-driven approach for fusion of several imaging modalities. Its main goal is to model the group data as a set of interpretable features (ICA components), each one characterizing a single, biophysically plausible form of variability. As shown in the right half of Fig. 1, each feature consists of a shared *subject loading*, which indicates which subjects have more or less of this feature, and the corresponding spatial pattern that is learned for each modality. It also adaptively allows a given feature to be completely absent from some modalities; this is especially useful for detecting artifacts, which are generally restricted to a single processing pipeline.

Linked ICA is implemented as described in detail in our earlier paper (Groves et al., 2011) with some minor improvements described in the following subsections. In this work we use the "flat" (as opposed to tensor) model configuration which means that while it assumes spatial alignment *within* each modality across subjects, there are no assumptions made about the spatial alignment *between* modalities. The lack of any assumptions about spatial alignment between modalities means that when we find changes that are spatially close to one another in different modalities, we can be confident that this is solely due to changes co-occurring in the same subjects across the different modalities, i.e. they are linked only by the shared subject loadings.

Correction for spatial smoothness

Linked ICA explicitly takes into account the spatial correlation of each modality, which helps ensure the correct balancing of information across modalities. In principle, this means that (for example) upsampling the data to a higher resolution (without changing its spatial frequency spectrum) does not add any additional weight to that modality. The effective number of independent measures was reduced using a data-driven degrees-of-freedom (DOF)/voxel estimate to perform what is known as "virtual decimation"; this is a necessary correction for spatial smoothness as the model would otherwise assume spatially-independent white noise. In our previous work we estimated this spatial smoothness using the RESEL estimates by Worsley et al. (1995). We now use a more flexible method based on the steepness of the eigenspectrum (Beckmann and Smith, 2004; Johnstone, 2001); the flatter the eigenspectrum, the more excess DOF there are in the data (i.e. spatial DOF >>#subjects). This is formalized by fitting the distribution of eigenvalues in the (null) multivariate normal.

This approach has several advantages over our previous method. Since we are attempting to estimate the noise smoothness, it is important to ignore the signal in the data which may be much smoother than the noise; our model assumes that these signals are low-rank, so they are concentrated in a small number of large eigenvalues. Similarly, preprocessing steps such as de-meaning can reduce the rank of the data, producing zero eigenvalues. As the null eigenspectrum only has two parameters (noise amplitude and spatial degrees-of-freedom), only two points on the eigenspectrum plot are needed for a fit; we use the 25th and 75th percentiles of the eigenvalues for this fit. As shown in Fig. 2a, this avoids these signals and preprocessing artifacts and focuses on the central part of the eigenspectrum that will be dominated by the noise floor. A secondary benefit is that this works without requiring information about spatial adjacency of measurements, which simplifies working with surface data (or potentially non-spatial data).

Dimensionality selection

We ran the linked ICA decomposition at two dimensionalities, $L = 50$ and $L = 100$. Based on previous experience, we have found that this data is rich enough to estimate at least 50 components robustly, and the upper limit of $L = 100$ was selected due to computational limitations; increasing dimensionality much beyond this slowed down the computation considerably and expanded memory usage beyond an acceptable level. Ultimately the "optimal" dimensionality depends on the quality of the data and the detail desired from the decomposition. Linked ICA is capable of eliminating unneeded components (using Bayesian model order selection), but in all cases it kept all components. Unless stated otherwise, the results in this paper are from the 100-dimensional Linked ICA decomposition; the 50-dimensional run is only used in the evaluation of component robustness.

Subject-wise noise estimation and missing data

A common problem in ICA decompositions is that a significant fraction of the ICA components will be dedicated to explaining single-subject oddities or artifacts. We reduce the impact of this by estimating the noise on a subject-by-subject basis within each modality; thus a single outlier scan that is poorly explained by the current components will be automatically inferred to have more additive noise, and thus less weight in determining the updates for these components. This is conceptually similar to the approach used in FSL's approach to group linear regression (Woolrich et al., 2004). This almost completely solves the problem of single-subject components, although outlier patterns that are consistent across several subjects may still be assigned a component (as will be shown in the results section). This same mechanism makes it trivial to deal with missing scans, by simply fixing the noise precision to zero, thus giving the missing data no weight in the ICA decomposition.

Post-hoc analysis and visualization

Each of the linked ICA components is defined by a subject-course (a vector of subject weights, one scalar value per subject), as well as *each* modality's corresponding spatial pattern.

As a final step, the subject-courses were fit back into the original high-resolution datasets without changing the same subject-courses or mixture models (spatial map histograms). The subject means were also reintroduced back into the appropriate components. This is conceptually similar to regressing the subject-courses back into the original data, as is done as a final step in standard probabilistic ICA (PICA) such as MELODIC (Beckmann and Smith, 2004; Johnstone, 2001).

The spatial patterns were converted to *pseudo-Z-statistics* by accounting for the scaling of the variables and the SNR in that modality, and are thresholded at $z = 3$ for visualization. Note that the model used by Linked ICA (and Bayesian ICA in general) is quite different from the mixture modeling approach used by PICA. The main reason for this is that Bayesian ICA uses an explicit spherical noise model during the decomposition. We therefore use this inferred noise level

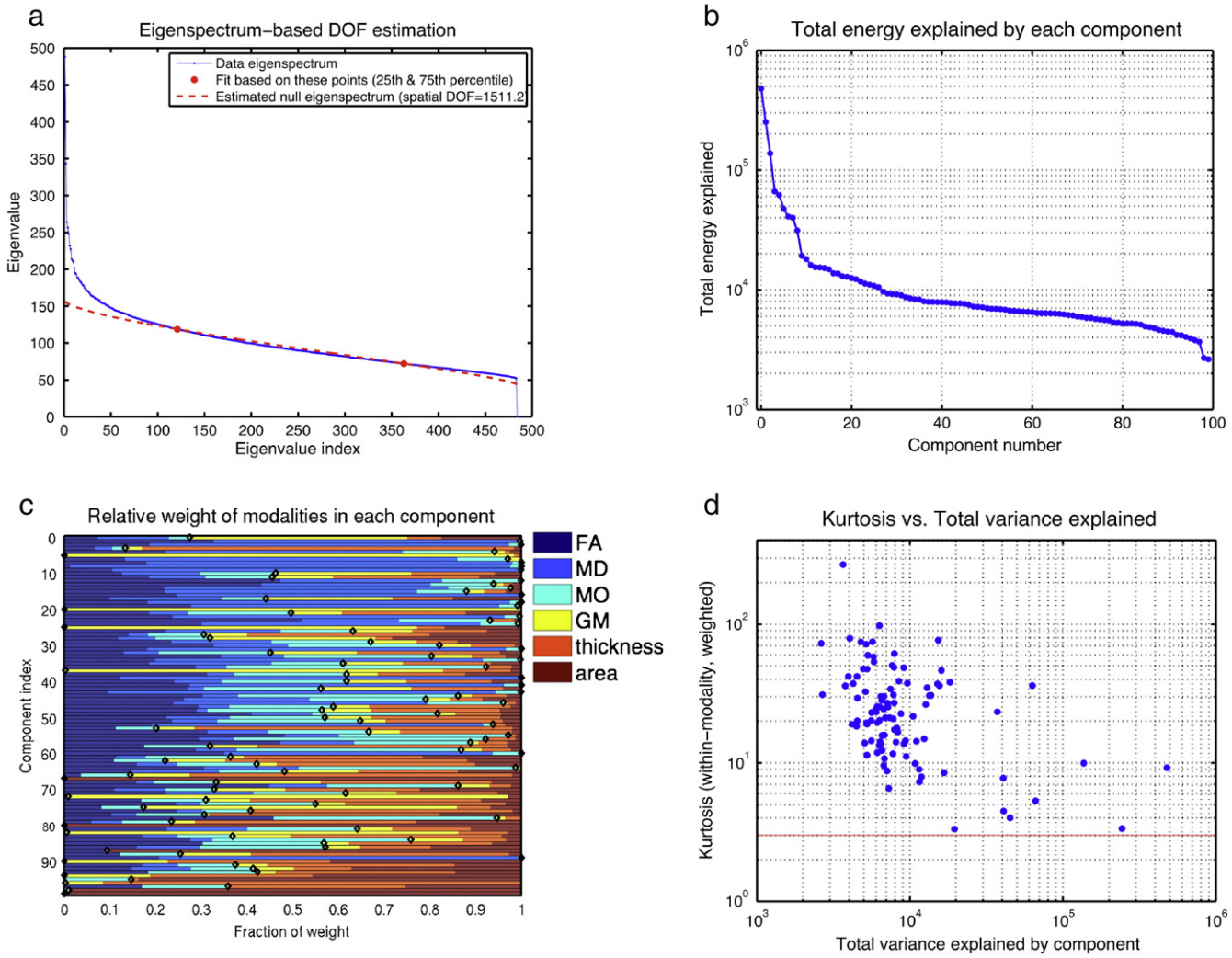

Fig. 2. (a) An example of the eigenspectrum-based DOF estimation, which corrects for spatial smoothness in the raw data — in this case, the tensor-mode (MO) modality. (b–d) Show summary information for the components found by Linked ICA: (b) Shows the total variance explained by each component. (c) The relative contributions of the six white and gray matter modalities to each of the 100 components. The diamond separates white matter (left side) from gray matter modalities (right side). Note how most components have a significant contribution from both GM and WM modalities. (d) Shows that the largest-variance components tend to be most Gaussian (lower kurtosis), while the low-variance components have the sparsest spatial maps (high kurtosis).

to determine Z-statistics rather than fitting it post-hoc from the spatial map histograms.

The components are sorted in terms of total explained variance (scaled by the noise in each modality). The sign of the components is arbitrary (in the sense that for a given component the subject-course *and* all spatial maps can be inverted in sign to give the same fit to the data).

The subject-courses were correlated (and plotted) against age, sex, and ICV (Buckner et al., 2004). By "significantly correlated" in this paper, we mean that a significant ($p<.05$) amount of the variance of the *subject-course* is explained by regressing against a particular *subject variable*. This is multiple-comparison corrected for the number of components (100) and also for the two tails. We did not initially correct for the number of subject variables evaluated, but all significant correlations reported in this paper also survived this further correction.

Age is a special case because we expect strong non-linear changes with age, so we fit the age curves using a nonparametric smoothing spline method with automatically-selected smoothing parameter, which Fjell et al. (2010) have shown avoids some of the biases of the more traditional polynomial fits. Furthermore, in components where the subject loading changes relatively monotonically with age, the dependent and independent variables can be flipped and the relationship can be used to *predict* subject age, given the loading. These fitting accuracies will also be reported.

Splitting component information by modality

In each iteration of Linked ICA, the shared subject-courses are updated by combining information from all modalities, averaged together using weights that reflect each modality's confidence in its estimates. These precision contributions appear naturally in the Variational Bayes updates for the shared subject loadings, and are an important summary measure for understanding which modalities are strongly driving a component, and which modalities are simply detecting weaker traces of the same signal. The precision contribution depends directly on the strength of the component in that modality as well as the modality's noise levels and spatial-smoothness.

It is possible to separate the contributions from the six modalities, yielding six slightly different estimates of the particular subject-course. This is useful for assessing which modalities are driving the significant correlation between a component's subject-course and a particular subject variable.

Assessing reliability of components

The reliability of the decomposition is assessed in several ways. We first examine the consistency of the components found by the two different dimensionalities, L = 50 and L = 100; ideally many of the components will be very similar, with L = 100 showing additional components and sometimes splitting of the L = 50 components into finer subdivisions. We also examined the sensitivity to initialization, by initializing the method using PCA or random vectors.

We also examine split-half reliability of the decompositions. Reliability scores are calculated by split-half resampling of the data (Groppe et al., 2009). We split the data set into two independent sets of subjects, and the Linked ICA inference is performed separately for each set; no data are shared. After this, the set of components from each split is "greedily" paired (i.e. taking highest-correlated match first) to the set of full-data components, by using the overlapping parts of the subject-courses. This produces triplets of components (half1, full, half2). The reliability is then assessed per-modality by looking at the correlation of the spatial maps in that modality, between half1 and half2 (the spatial map of the full data set is not used for scoring). Significance is assessed by comparing these spatial map correlations to the null distribution of correlations between all other possible pairs of half1 and half2 components (i.e. those that are not part of the same triplet). This null distribution corresponds to the situation where the decomposition method does not detect any consistent structure in the data. For comparison, we also evaluate the reproducibility of components found by a simple spatially-concatenated PCA on the same data set.

Results

Fig. 2b depicts the relative amount of the total variance explained by each component, sorted in a descending order for the 100-dimensional decomposition. This shows a relatively clear elbow at around component #10, possible signifying a shift in the qualitative features of the components between the first 10 and the last 90 components (see the Global components subsection below).

Fig. 2c shows the relative modality loadings in each component; it is clear that while no one modality was dominant overall, many of the components were modality-specific (one modality strongly dominates). Several components were dominated entirely by VBM, others by DTI, in particular MD. These represent types of inter-subject variability that are only present in one modality or one group of related modalities. However, a significant fraction of components show a relatively equal balance between DTI and morphometric modalities, indicating patterns of variability that were observable in both diffusion and structural scans. The general GM–WM split was made visible by the diamond symbol in each row, which separates the WM signal (to the left of the diamond) from the GM signal (to the right).

Neuroanatomical results

Global components

Linked ICA picks out a number of high-variance "global" components that have fairly Gaussian spatial maps and coexist with the remaining "local" components, despite having a significant overlap with them (Fig. 2d). We identified several of these global components as having a very high correlation with age. The two strongest examples are shown in Figs. 3a and b.

Fig. 3a (component #0) describes a set of inter-subject differences covering all modalities, and shows a strong correlation with age, indicating that this component describes age-related differences. Using the spline-fitting approach and leave-one-out cross-validation, the subject courses from this component "predicted" subject age with a standard deviation of 5.9 years overall ($r = 0.97$), improving to 2.4 years in the 8–16 year old subjects. We did not attempt to predict age from the other components as the subject-courses do not show a monotonic change with age.

The slope is significantly ($z = -11$, $p < 10^{-27}$) steeper in the developmental (8–25 years) compared to later phases (>25 years), indicating that the changes occur more quickly in the young subjects.

With some notable regional exceptions, as age increases MD is primarily decreasing along with decreasing FA. An opposite pattern is observed in areas around the brain stem and thalamus for FA, and parts of the genu and the external capsule for MD. MO shows a heterogeneous pattern of decreases in frontal areas including the forceps minor and increases in regions including parts of parietal lobe. The GM measures are in general consistent with concurrent reductions in volume (VBM), thickness and area from the earliest age sampled.

Fig. 3b (component #4) shows a different age-related pattern, closer to a U-shaped trajectory. In this case the subject-course is strongly non-linear, with nearly no change up to age 50–60 and then changing at an increasing rate. The largest observed changes are the widespread increased MD, especially in major tracts such as the corpus callosum. FA and MO are generally seen to be decreasing, especially in peripheral tracts. Interestingly, one bilateral region encompassing the superior longitudinal fasciculi shows the opposite effect, with FA and MO both increasing.

We also observed gray matter involvement in this component, primarily in terms of reduced gray matter volume as measured by VBM. The reduction appears to focus on subcortical structures, including the putamen, amygdala, caudate, accumbens and thalamus. Parts of the cerebellum also show volume reductions, as does the medial occipital cortex.

Fig. 3c (component #1) shows fairly global increases in MD along with global reductions in FA. It has a similar U-shaped age profile to #4 but appears to be "noisier", suggesting that there are other factors driving these global white matter changes that are not as clearly related to age. There are no gray-matter changes in this component. It is also interesting to note that although components #1 and #4 have very similar U-shaped age-profiles (fits are correlated with $r = 0.96$), the subject loadings have a much lower correlation ($r = 0.24$) that is no longer significant when the age fit is regressed out ($r = -0.08$). This implies that the age-related changes described by these two components have very different patterns of inter-subject variability.

Fig. 4 shows two additional strong multimodal age components. Component #3 showed a very strong correlation with ICV ($r = 0.88$, $z = 26.5$). Since ICV and sex are highly correlated, this component also has a strong correlation with sex ($r = 0.63$, $z = 15.6$); and indeed both regressors are still significant in a multiple regression, but ICV is clearly the dominant effect (ICV $z = 22.3$, sex $z = 5.9$). The spatial maps show strong global increases in surface area corresponding with a globally increased ICV, as well as some scattered changes in cortical thickness. There are also some weak diffusion differences associated with this component, which appear to focus mainly on the periphery of the skeleton, suggestive of predominantly partial volume effects. There is minimal VBM involvement in this component, but it is worth noting that this measure would only be sensitive to *relative* gray matter volume changes, as the global scaling of each brain is absorbed by the initial affine registration and not included in the VBM measure.

Fig. 4b shows component #17, which is primarily related to sex differences ($r = 0.28$, $z = 6.2$), which becomes stronger once ICV is regressed out ($r = 0.37$, $z = 8.5$), suggesting that this component explains sex differences over and above ICV. In a multiple regression, both are significant (sex $z = 8.5$, ICV $z = 5.8$). This component corresponds to a pattern of specific increases in cortical thickness in the cingulate along with reduced gray matter volume in parietal regions and higher MD in peripheral white matter in males compared to females. Notably, this component has a U-shaped age profile, with the sex differences strongest in the middle age range.

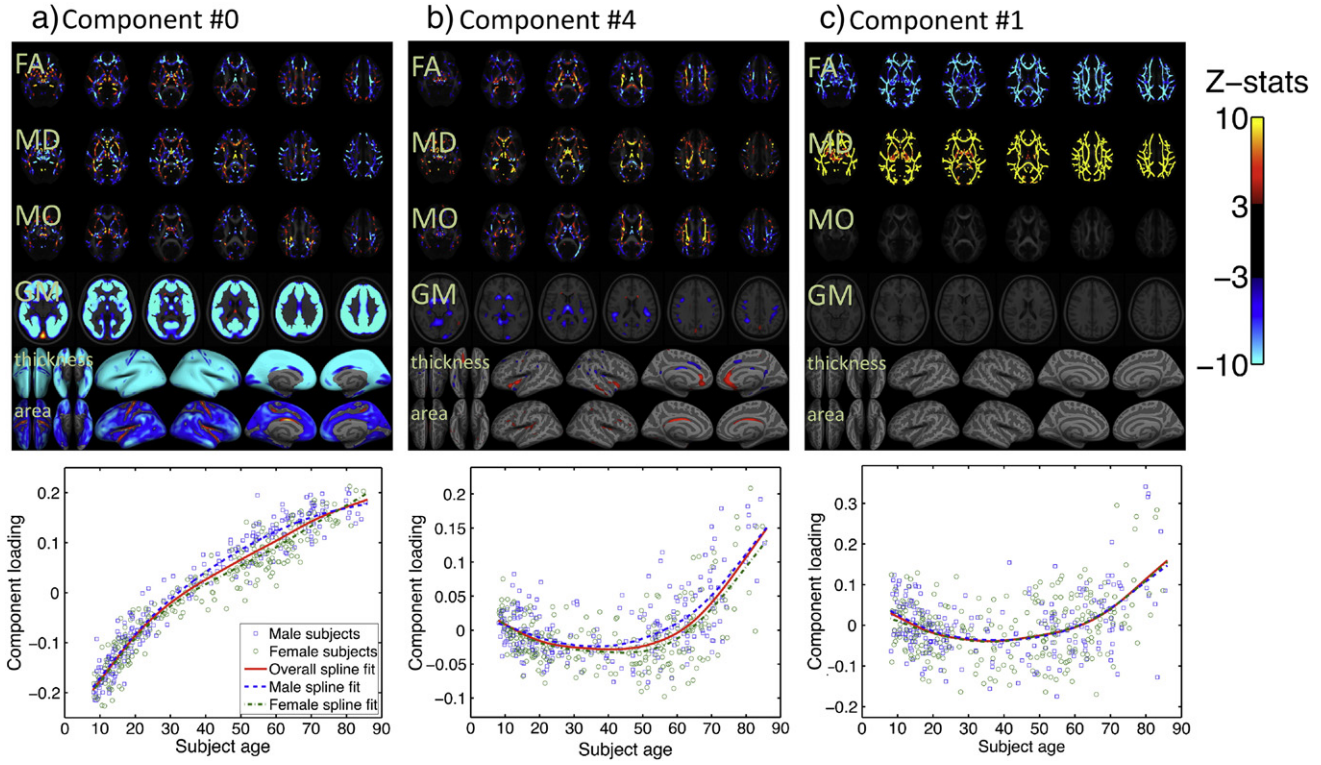

Fig. 3. Three of the strongest components show age-related effects. (a) Component #0 pulls out a multimodal set of brain differences that is dominated by cortical thinning and VBM decreases, along with a strong pattern of changes in FA, MD, and MO. This component is very highly correlated with age ($r = 0.95$), and has steeper increases in development. (b) Component #4 has a U-shaped age profile that especially emphasizes differences related to aging, with patterns of white matter change and localized GM decreases. (c) Component #1 shows a similar age profile, with widespread MD increases and FA decreases in old age.

Assessing the contributions of each modality to global components

Although Linked ICA is built on the assumption that a shared subject-course can describe a component across several modalities, it is also possible to calculate the individual contributions of each modality to the subject-course (see Appendix A). Conceptually, this provides an estimate of how well the estimated subject-course would match the external variable if only one of the modalities was available. This shows us whether the modalities that dominate the component (by having large "precision contributions" in the VB updates) are actually those that are most directly related to the relevant subject variable (e.g. age). Fig. 5 shows a few examples of this. The right bars are the correlation between the subject-courses and a nonlinear spline fit of age (or ICV) to these subject-courses. Alternatively, a regular correlation can be used, which assumes a linear relationship between subject-course and age. This provides very similar results. The bars on the left (component weights) relate to the *signal-to-noise* ratio that shows the overall strength of that component's spatial map in each modality, while the right bars relate to the *contrast-to-noise* ratio of the subject variable of interest.

In Fig. 5a, we see that cortical thickness (CT) and gray matter volume (GM) dominate component #0. However, each other modality, taken alone, is still able to predict age with high accuracy ($r > 0.93$). The combined subject-course (represented by the "all" bar) is closer to the true age than any individual modality.

In Fig. 5b, we see that the gray matter pattern in component #4 is slightly more informative about subject age than either FA or MO, despite having a much lower weight in the decomposition.

Fig. 5c shows component #1, which is dominated by FA and MD. In fact, including the other modalities makes the age prediction *worse*. Fig. 5d shows the relationship between component #3 and ICV; in this case arealization (CA) is dominant, but including the other modalities does improve the prediction slightly.

Structured noise components

Fig. 6a shows component #6, which is related to scanner software upgrades partway through the data collection (i.e. Siemens software Syngo v. 13, 15 and 17). The asymmetrical pattern is strongest in MD and more weakly present in FA and MO. Fig. 6b shows a similar artifact in component #2 which is also related to scanner version.

Fig. 6c depicts component #14, which shows apparent WM differences around the lateral ventricles. The corresponding age profile shows that the subject loading is strong only in a small number of elderly subjects. These subjects have large ventricles, and so this component could partly reflect misregistration of these areas due to these subjects' WM tracts being too far removed from the TBSS template, or alternatively subtle periventricular vascular insults.

Component #52, in Fig. 6d, also only affects a small number of old subjects, but an entirely different subset. In addition to DTI differences these also show deep gray matter volume differences in the caudate and part of the putamen, along with nearby WM changes. A closer examination of the T1-weighted volumes of these few subjects suggested noticeable signal alterations in the areas of gray and white matter changes, and this component isolates the effects of these.

Localized related changes in gray and white matter

Fig. 7 shows a few of the substantial number of components describing anatomically-related, spatially focal GM–WM differences. These mostly indicate relationships between thickness and volume and the integrity (FA and MO) of the corresponding WM pathway. None of these components correlated significantly with age, sex or ICV.

Fig. 7a (component #11) shows a correlated pattern of FA and MO in the posterior part of the forceps major and likely the optical

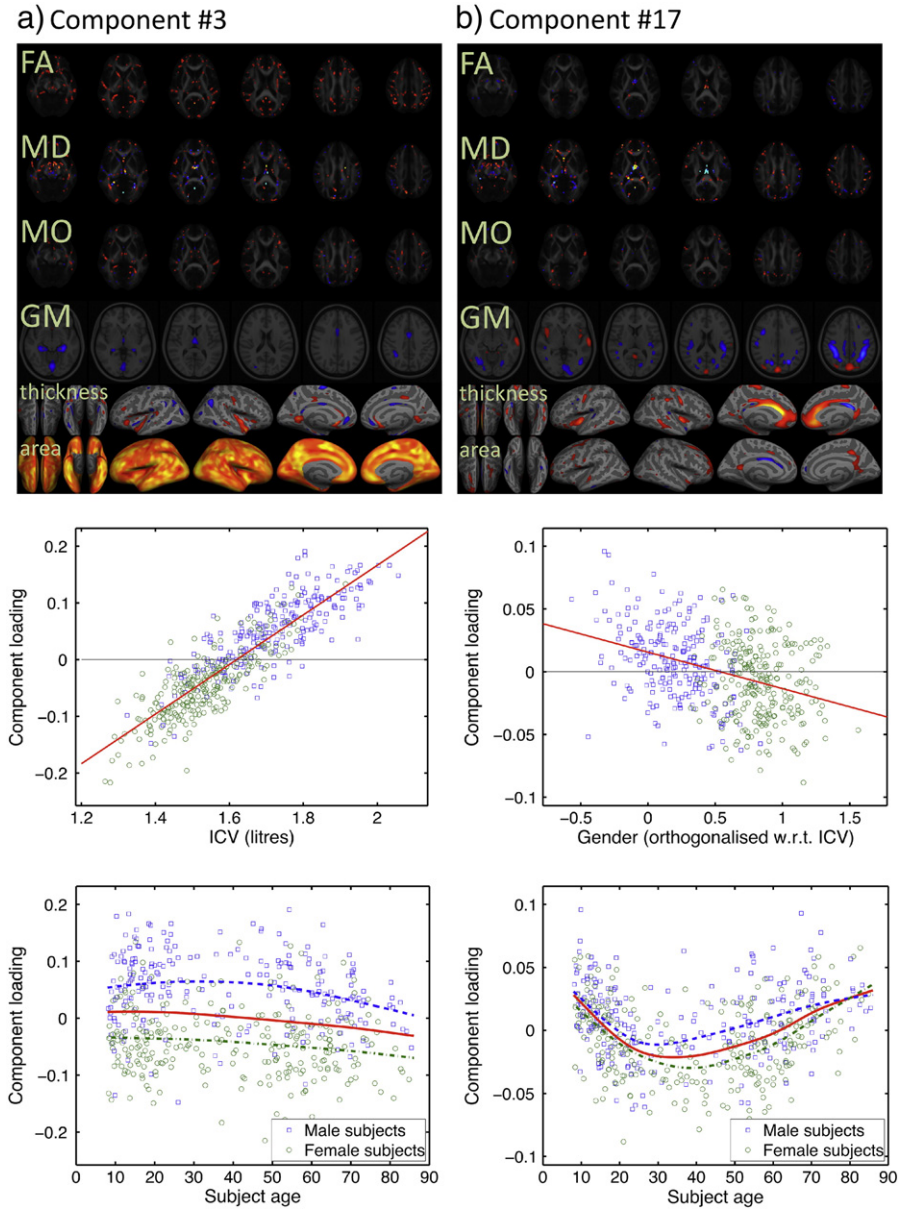

Fig. 4. Global components that are related to intracranial volume (ICV) and gender. (a) Component #3 is very strongly associated with ICV. (b) Component #17 is highly correlated with sex, especially after ICV is included as a confound regressor. This component also changes with age.

radiation. This would normally be associated with higher general integrity of this tract, although the associated higher MD suggests this may be more complex. Alongside this tract there is lower surface area in the medial visual cortex, with the more anterior part showing higher cortical thickness and the posterior part showing lower GM volume (VBM).

Fig. 7b (component #38) also shows higher FA and MO in bilateral tracts comprising the visual pathways feeding into the occipital cortices. This WM pattern is associated with lower gray matter volume in medial and lateral occipital areas, as well as slightly increased volume in a visual area that appears to lie at the terminus of the involved white matter pathways.

Figs. 7c–d depicts two other components of GM–WM differences in the cingulate sulcus. Unlike previous components, these are fully lateralized; component #42 is almost entirely in the left hemisphere while #48 is almost a mirror-image of it. Large increases in cortical area are adjacent to large decreases, with VBM and thickness following a similar pattern. The FA and MO changes are immediately adjacent to this, showing a spatially very similar pattern of increases and decreases.

Relationship between the DTI indices

Fig. 8 shows a number of components describing modes of variation primarily in the white matter; in some cases there are also weak GM thickness/volume differences in the same areas (not shown). The components in Fig. 8 involve major fiber pathways and come in two distinct flavors: FA + MD in the largest white matter tracts and FA + MO in smaller tracts. In all cases these effects appear bilaterally in local regions of the TBSS skeleton.

Figs. 8a–c show three components that isolate the posterior corpus callosum, anterior corpus callosum and forceps major, respectively. These show large increases in MD along with reductions in FA, which is a pattern generally consistent with integrity of major tracts with minimal crossing fibers (Douaud et al., 2011). There is very little involvement of MO, possibly because this measure is less sensitive than MD and FA when there is no clear secondary diffusion direction.

Several examples of components with combined FA and MO changes in the smaller tracts are shown in Figs. 8d–g. In these components there is no significant MD involvement, but spatially overlapping

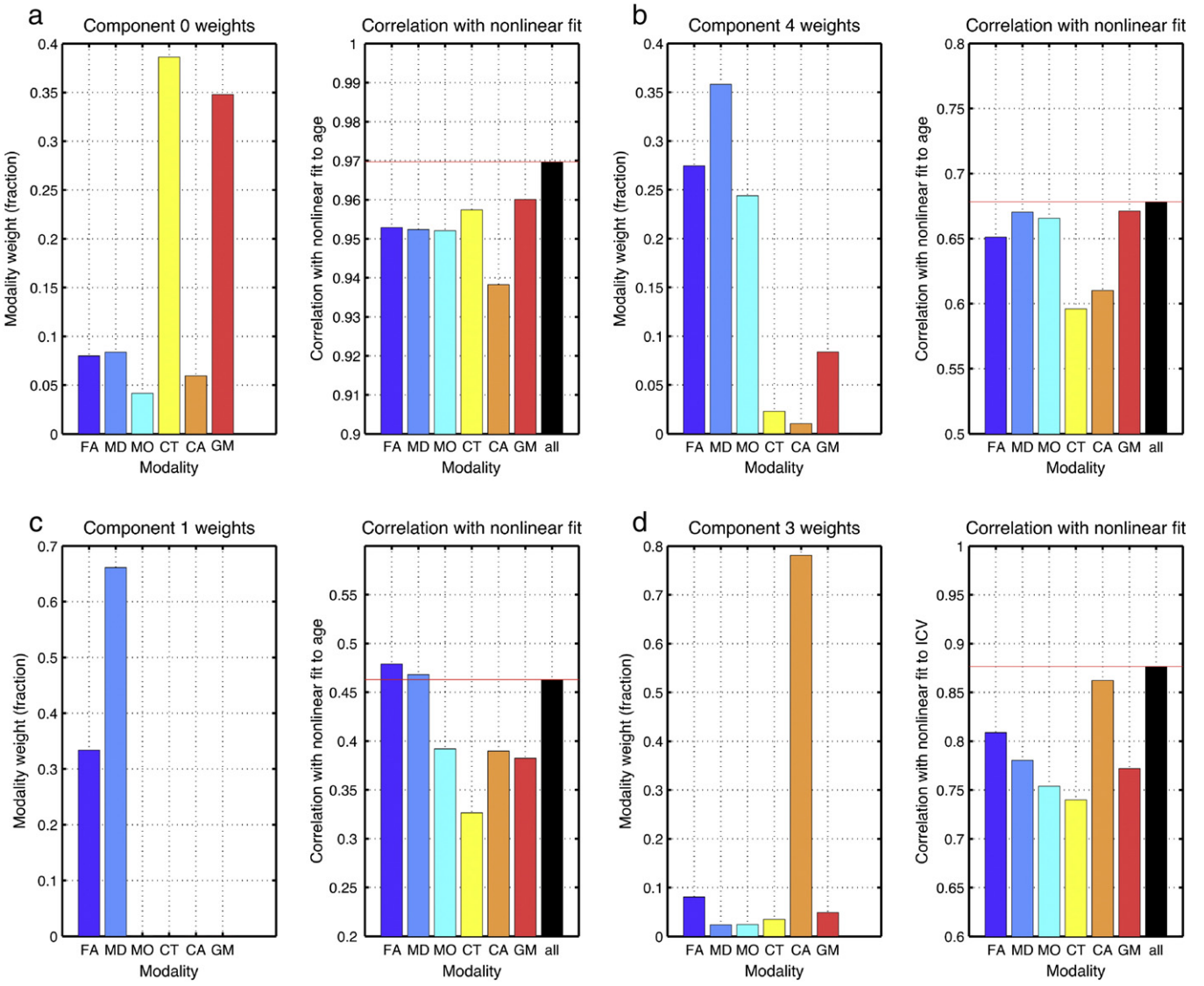

Fig. 5. It is possible to split contributions of each modality to a component. (a) Component #0 and (b) component #4 both predict age most accurately when all modalities are used. (c) Component #1 is driven by MD and FA, but FA is actually better than the combined fit; this suggests that the MD pattern is responding to something other than just age-related differences. (d) Component #3 predicts ICV most accurately when all modalities are used, even though cortical-area is by far the most heavily weighted.

increases in FA and MO. These appear to be showing change in the integrity of the larger fiber tract, while a smaller crossing tract is unaffected. The relative strength of these modalities varies somewhat between components, and even in different regions of the same component. None of these components correlate significantly with age, sex, or ICV.

Methodological results

Sensitivity to initialization

We re-ran the linked-ICA decomposition using a different initialization for the subject-courses, **H**; instead of a PCA decomposition on the smoothness-weighted, concatenated data we used Gaussian random noise and initialized the spatial maps by multiple-regression of these random subject-courses into the data.

The subject-courses resulting from each Linked ICA run were greedily paired between both runs and the absolute correlations are plotted in blue in Fig. 9a. The concatenated spatial maps were similarly paired, and these correlations are shown as the green line. The plot shows that the inferred components are nearly identical under each approach, with the median component pair having subject-courses that are $r = 0.98$ correlated and spatial maps that are $r = 0.95$ correlated.

We repeated this procedure, but instead of initializing with random subject-courses we initialized using covariance-matched Gaussian noise, where $N(0,1)$ noise has been post-multiplied by the matrix square root of the cross-subject covariance matrix. These results are very similar, and are shown with dashed lines in Fig. 9a.

Comparison of two different dimensionalities

Fig. 9b shows how well the components found by a lower-dimensionality decomposition ($L = 50$) have matching counterparts in the $L = 100$ decomposition that was used for our main results. This shows that most components are quite stable across this large increase in dimensionality, with a median spatial and temporal correlation over 0.7. This means that while some of the components split when there are more dimensions available, others are very similar and the new dimensions are filled with additional components that would have otherwise been treated as noise.

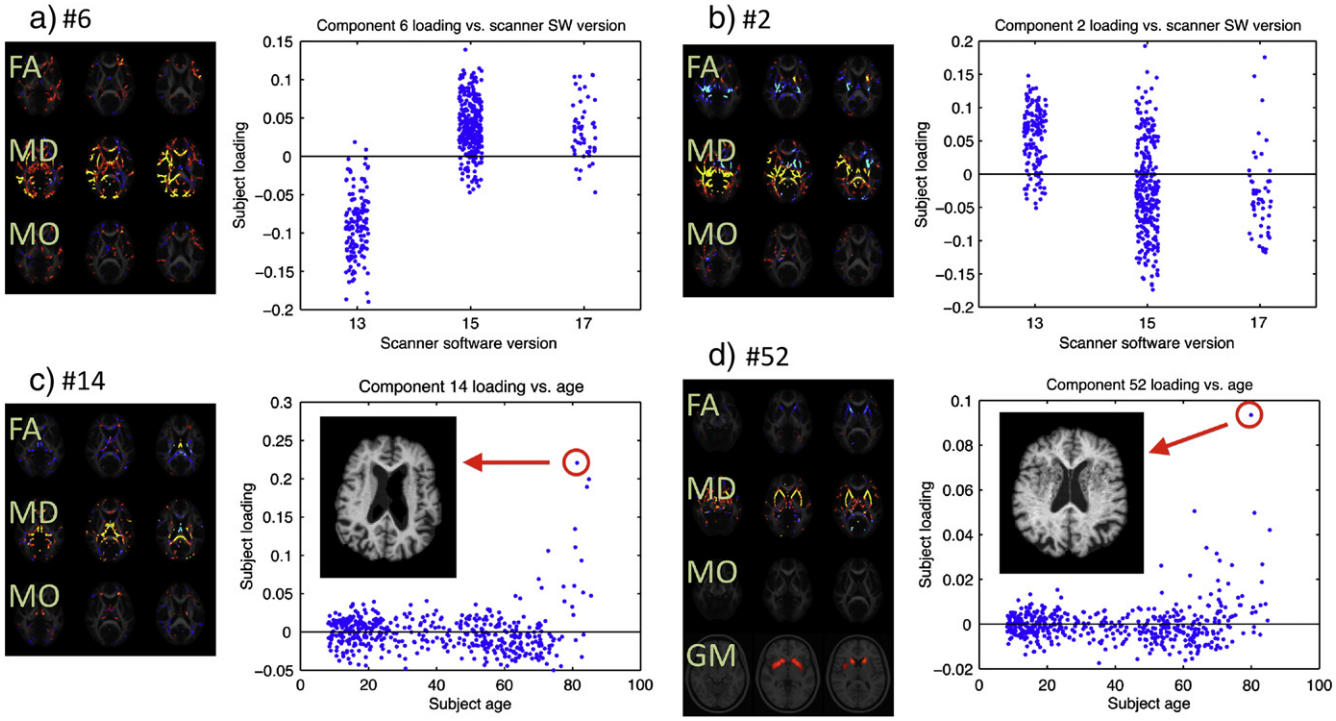

Fig. 6. Several artifactual components discovered by linked ICA. Modalities with no voxels above threshold have been omitted. (a–b) Asymmetrical patterns in the DTI modalities (especially MD) that are related to scanner software upgrades. (c) A registration artifact related to expansion of the lateral ventricles in some very old subjects. (d) A set of changes that are related to apparent subcortical and white matter calcification in several old subjects. The brain slices come from the highest-loaded subjects in components #14 and #52, to illustrate these problems.

Split-half reliability

Fig. 9c shows the split-half reliability of each component. Each component in the full decomposition is greedily paired with one component from each half-decomposition, using only subject-course correlations. The *spatial* correlation between each half-decomposition's spatial maps is used as a "reproducibility score". For comparison, the concatenated PCA produces much lower reproducibility score for most of its components. This demonstrates that our ICA-based approach is finding features that are more stable and reproducible than simple variance.

We can determine how many of these pairings are significant by building an empirical null distribution, based on the spatial correlation between all *unpaired* half-decomposition components. Using false discovery rate (FDR) thresholding ($\alpha = 0.05$), we find that 74 of the linked ICA components, but only 17 of the PCA components, are significantly reproducible (Fig. 9d).

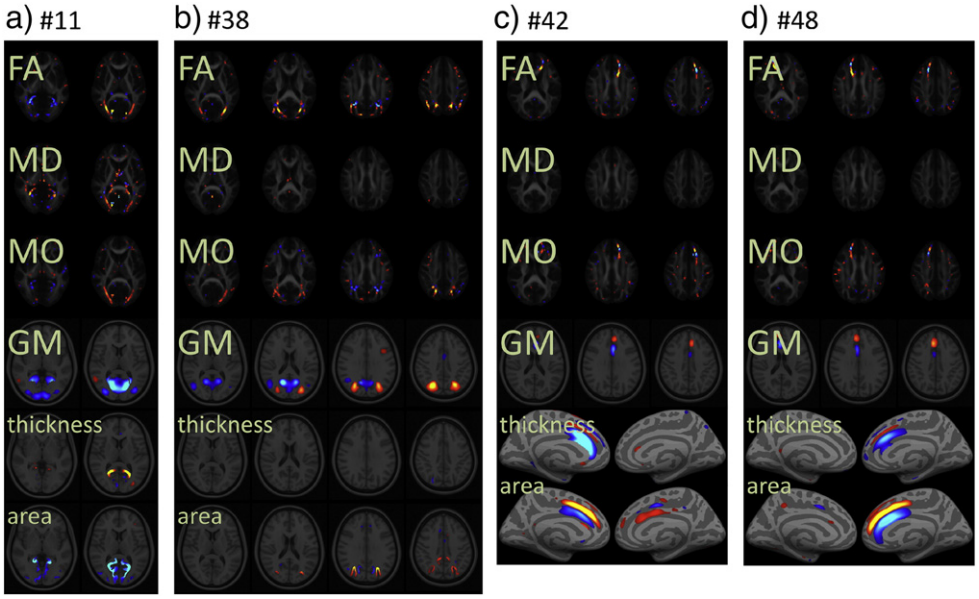

Fig. 7. Four components showing related differences in gray matter and white matter. (a) shows that increased FA and MO in the optic radiation is associated with reduced GM area and volume in parts of medial V1/V2 — but also increased gray matter thickness in the same area. (b) shows that increased FA and MO in a bilateral pair of tracts comprising the visual pathways feeding into the occipital cortices is associated with GM area and volume changes in the related intraparietal region. (c–d) show related GM–WM changes in the left and right cingulate, respectively.

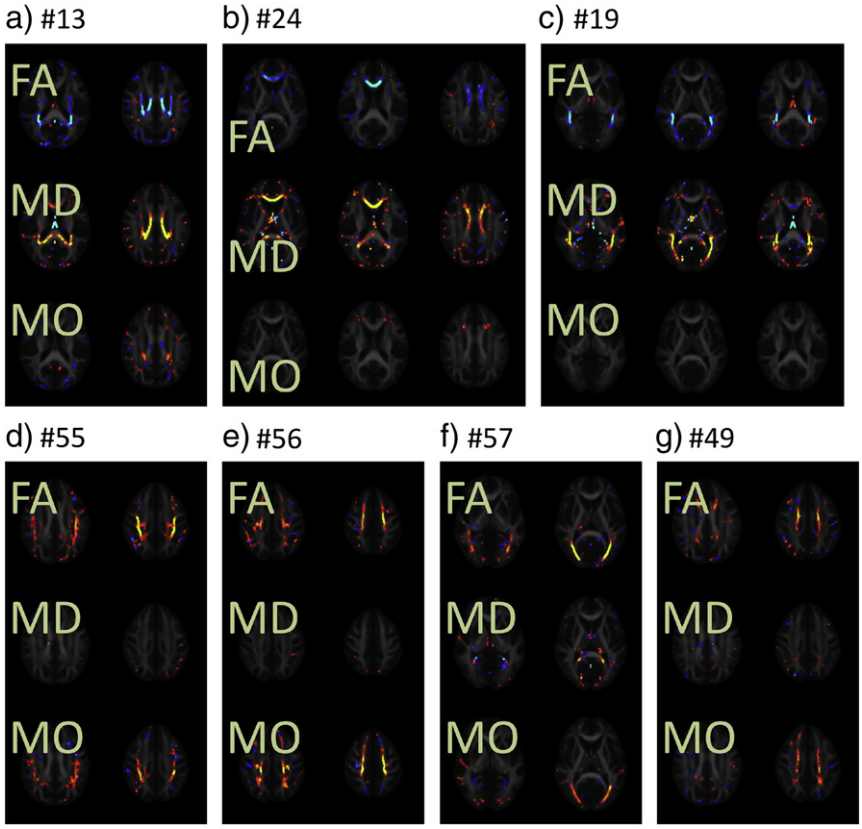

Fig. 8. DTI components showing related TBSS changes in FA, MD, and MO. (a–c) show three white matter components with overlapping, opposite-sign differences in MD and FA. (d–g) show four white matter components with overlapping, same-sign FA and MO changes.

Discussion

We have demonstrated that Linked ICA decomposes a large multimodal data set into components that fall into a number of distinct categories based on spatial pattern and the modalities that are involved. Broadly, we have reported a relatively small number of "global" components that match up well with age, sex, and ICV. Many more "local" components were found reflecting spatially distinct, generally bilateral, and usually multimodal variability. Lastly, several components reflecting structured noise or imaging and/or analysis artifacts were found. We will discuss the various modes of variability in more detail below.

Global components

The global components contain concentrated information about individual subject variables and show very strong relationships with age and ICV. These components have spatial histograms that are close to Gaussian, and are driven more by variance than non-Gaussianity. It is worth noting that this is quite unlike the behavior of standard probabilistic ICA (Beckmann and Smith, 2004; Johnstone, 2001), which uses a prewhitening stage to totally remove the influence of variance. Linked ICA is instead based on a generative model of the original, unwhitened data, which makes it more sensitive to large-variance signals and therefore combines some PCA-like behavior with ICA-like behavior. For this reason we suspect that these global components may change in form if the composition of the population changes; the age-related components we found may only be the most prominent ones because these subjects are spread over a wide age range. There is a sharp transition, with components 0–9 explaining large amounts of variance and being spatially global, while 10–99 explain a relatively consistent amount of variance each but are much more focal.

The subject courses of the largest component showed very high sensitivity to age, and prediction analysis using smoothing splines and leave-one-out cross-validation revealed high prediction accuracy across the full age-span ($r = .97$), and even higher when analyzing a cohort of the youngest sample. This prediction accuracy compares favorably with previous work on predicting age using functional network data (Dosenbach et al., 2010) and is comparable to previous results using VBM alone (Franke et al., 2010). It is important to remember that linked ICA is an unsupervised learning process that found this component without any knowledge of the age of the subjects, and that instead of regressing out confounds explicitly, this approach automatically removes structured confounds by isolating them in other components. We did not attempt to predict age from the other components as the subject-courses do not show a monotonic change with age. Bearing in mind that linked ICA is totally naive to subject age, we find the high age sensitivity of some extracted components encouraging and supportive of our assumption that the multimodal decomposition extracts biologically meaningful information. This is further supported by our finding that age prediction accuracy was higher when using our multimodal component as compared to using the same information from each of the modalities individually.

Some of the components, when considered alone, appear to complement results from simpler analysis methods; for example, the age-related changes in component #0 suggest relatively linear effects of age across the lifespan, but demonstrate novel complex spatial patterns of both positive and negative associations with age. This observation extends previous findings of age-related differences in the DTI indices (Westlye et al., 2010b), and demonstrates the power of the current method in detecting patterns of age-related brain differences across subjects. Crucially, the spatial maps found by linked ICA are effectively a *multiple* linear regression, where the data set

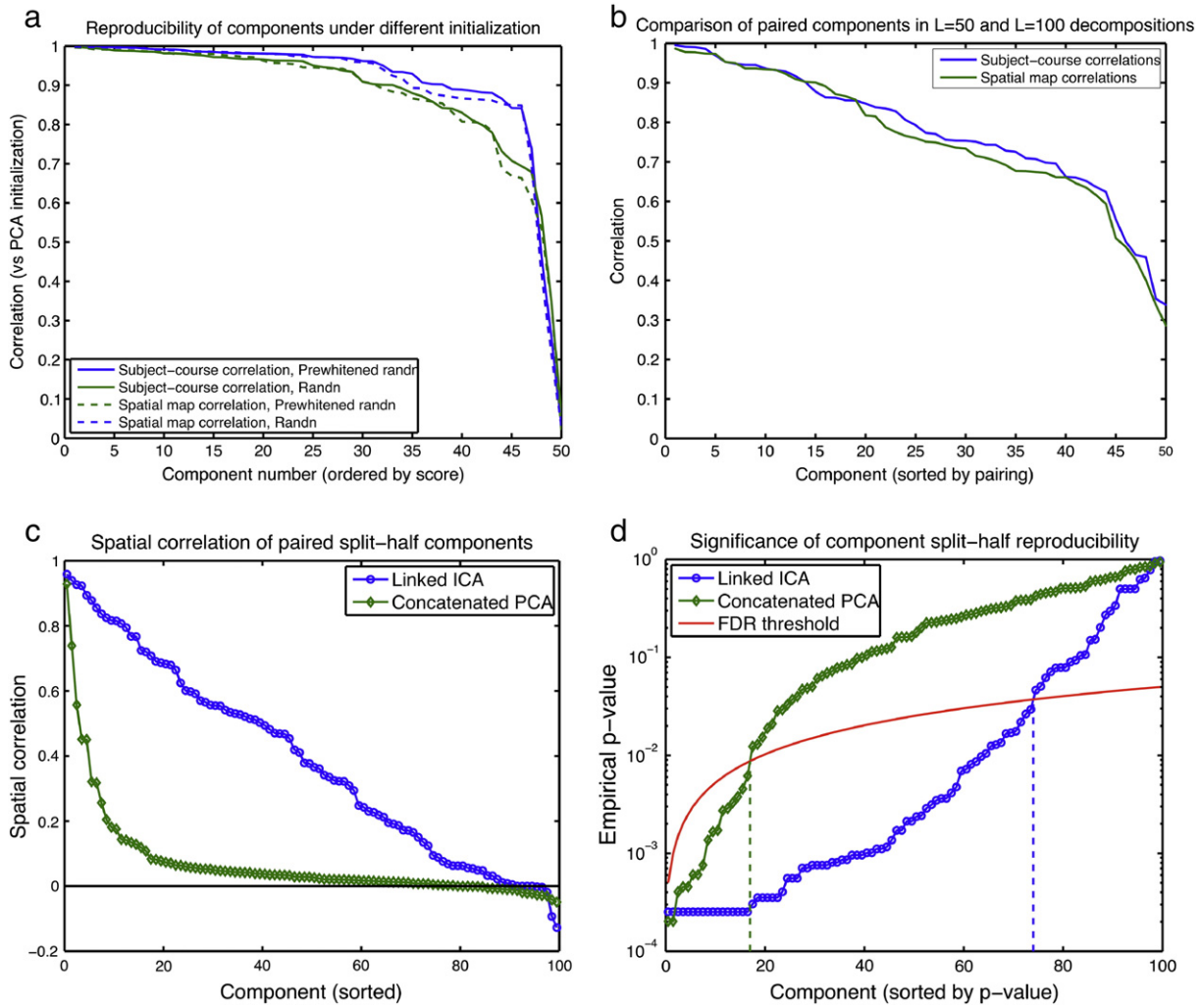

Fig. 9. Assessments of reproducibility, based on pairing components across different decompositions. (a) Initialization using PCA vs. random initializations. (b) Comparison of two model orders: L = 50 and L = 100. Many of the components are highly similar, with some changing due to e.g. splitting. (c) Split-half reliability results, showing the spatial correlations between components paired using only subject-course information. (d) Using an empirical null distribution and false discovery rate (FDR) thresholding, we assess the significance of these pairings.

is explained by the combination of all components; so age-related changes that appear absent in the main component could be better explained in another component that is also correlated with age. For example, the global cortical surface area is dominated by age-related decreases (in #0) and ICV-related increases (in #3).

The patterns found in other global components are in general agreement with previous regression-based studies in earlier analyses of overlapping samples (Tamnes et al., 2010; Westlye et al., 2010b). The main nonlinear age component (#4) agrees with the nonlinear trends in gray matter volume found by Ziegler et al. (2011) and in a multi-sample study (Walhovd et al., 2011). Similarly, Westlye et al. (2010b) showed differences in white matter volume and DTI indices that find similar quadratic age profiles with increasing FA and decreasing MD until the 4th decade, followed by declining FA and increasing MD, which is also supported by independent DTI studies (Hasan et al., 2007, 2009; Imperati et al., 2011; Lebel et al., 2012).

The opposing white matter changes shown in component #4, where FA and MO *increases* are associated with aging in some tracts, are believed to be areas of crossing fibers, since increases in diffusion anisotropy can be caused by the weaker tract degenerating while the stronger tract is preserved. A very similar pattern of increasing FA and MO was recently reported in degeneration related to mild cognitive impairment and Alzheimer's disease, and further corroborated by tractography analysis indicating that this specific pattern may be explained by degeneration of the association pathways and a relative preservation of the motor pathways (Douaud et al., 2011).

Structured noise components

We found several strong diffusion-only components that were clearly related to imaging artifacts with consistent spatial patterns across subjects. One of these was clearly related to a minor software upgrade, but we did not find an explanation for the other components. A significant benefit of an exploratory approach like ICA is that this can find structured noise in the data by modeling the spatial structure; this has long been used on FMRI and electroencephalography (EEG) data. Isolating artifacts as specific components prevents them from contaminating the remaining signals, and may increase sensitivity to correlation against subject variables. Our approach is particularly powerful when artifact components are present in several related modalities. For example, the diffusion artifacts are strongest in MD, which makes it easy to detect and characterize the artifacts and estimate their strength in each subject, so that they can be removed from FA and MO more cleanly that would be possible otherwise.

Local components

The local components that coexist with the global components are driven by non-Gaussianity and spatial sparsity in particular. We find that nearly all of the non-artifact components involve all modalities at some level, and some of them are very evenly balanced, showing corresponding differences in gray and white matter. To what degree these explain functionally-relevant variation in brain structure remains an open question. The specificity of the local DTI components (Fig. 8) may reflect modes of variation directly or indirectly modulated by experience and use-dependent shaping of cortical and subcortical networks (Zatorre et al., 2012). Future linking between the hierarchical organization of the components may substantially inform systems-based approaches to cognitive and clinical neuroscience. We are currently designing further analyses where we explicitly aim to map the neurobehavioral and genetic associations onto the multimodal components. We are also working on implementing linked ICA on longitudinal datasets in order to delineate the individual age-trajectories in more detail.

White matter components

In white matter, FA appears prominently in all local components of variation. However, for large tracts, FA and MD show predominantly negative correlations while MO is unchanged; and for smaller tracts, FA and MO show positive correlations as MD remains constant. The strong involvement of FA supports that FA is a sensitive general imaging marker for WM differences, but it is important to note that in some components the MO or MD differences were more significant than FA. It also suggests that a linear combination of these measures could be more sensitive for detecting certain changes than any single measure alone. Indeed, a variety of measures are used in the literature to gain sensitivity, such as the first eigenvalue of the diffusion tensor (L1) and radial diffusivity (RD). In the present study, we only included FA, MD and MO because these measures are mathematically orthogonal and should have independent noise from one another.

Similar to the focal GM components, the components dominated by DTI were also characterized by bilateral and anatomically and functionally related regions, including longitudinal (e.g. #56, reflecting parts of the superior longitudinal fasciculus) and commissural (e.g. #24, reflecting anterior parts of the corpus callosum) pathways. Some of these components are quite similar to those found by Li et al. (2011) using FastICA on FA alone. Our results shed further light on the meaning of these components by incorporating MD and MO in our model, which allows all three measures to drive the decomposition. The correlated pattern of differences in FA and MD/MO, respectively, is likely to be biologically informative. Increased FA along with decreased MD (e.g. #13, #19 and #24) may reflect biological variability related to the integrity of pathways with minimal crossing fibers, with very little involvement of MO, likely because there is no clear secondary diffusion direction and MO is saturated. Other DTI components (e.g. #49, #55, #56, #57) show spatially overlapping increases in FA and MO, but no MD involvement. This pattern may partly reflect variations in degree of crossing fibers, in particular differences in the integrity of the larger fiber tract, while a smaller crossing tract is unaffected. The strength of the modalities varies from component to component, and even in different regions of the same component, likely partly reflecting regional differences in degree of crossing fibers.

The findings of distinct and reasonable spatial maps reflecting regional cross-subject variability in WM microstructure inform the debate regarding regionally general versus specific involvement in modulating age- and behavioral intra-individual differences (Li et al., 2011; Lövdén et al., 2012; Penke et al., 2010; Wahl et al., 2010). The present decomposition supports a multidimensional perspective of WM organization, but further studies are needed to establish the true dimensionality and the hierarchical structure.

Multimodal involvement

Several components showed high loadings from both GM and WM modalities. While this pattern indicates shared variance across tissue types, we are not certain as to the underlying cause of these related differences. Little is known about the relative impact of common and specific modulators of gray and white brain tissue, respectively, but it is fair to assume a relatively large degree of shared influence across tissue classes, which could explain the various multi-modal components. Relating these components to differences in behavioral, environmental, and genetic measures is a promising avenue for future studies. Using the subject-course-splitting approach from Fig. 5, it may be possible to determine whether these effects are both directly due to genetic effects, or whether the genes more strongly drive one particular modality, and the other modalities show secondary effects.

The positive VBM and FA correlations in component #38 (Fig. 7b) appear in nearly the same location as the changes reported by Scholz et al. (2009) when subjects practiced juggling for several weeks; these were believed to be an area and tract related to visual motion tracking. It is unclear whether the observed components in the present analyses could be functional – for example, relating to experience-dependent differences related to these areas – or whether they are purely morphological. However, several components revealed an intriguing pattern across modalities, which may point to functional specificity.

We speculate that some of the cross-tissue components partly reflect functionally-relevant differences in gray matter and its corresponding white matter connections. One would generally expect these to show differences in the tracts that are connected to the gray matter rather than those that simply run past it, and the direction of the association should usually show larger gray matter volumes, thickness, and areas associated with improved white matter quality (generally increased FA and MO, reduced MD). For example, the multimodal spatial patterns of component #11 (Fig. 7a), showing decreased GM density and surface arealization and increased cortical thickness in the primary visual cortices along with increased FA, MD and MO in visual WM pathways, likely reflect structural differences in the visual system. A speculative explanation could be that if the primary visual cortices, including V1, cover smaller areas of the cortex – which is reflected in the decreased surface arealization – the white matter pathways feeding into it will appear “denser” and more anisotropic if it contains the same number of axons in a thinner tract. Furthermore, assuming that these smaller visual cortices contain the same number of neurons, this could explain the increased cortical thickness observed in this component. Although speculative, such interpretations may provide a functional basis for hypothesis generation in future studies. For example, individual differences in the size of the primary visual cortices have recently been linked to the subjective experience of object size and visual illusions (Schwarzkopf et al., 2011), and further to the peak frequency of visually induced oscillations in the gamma band as measured with magnetoencephalography, which is assumed to reflect important aspects of visual processing (Schwarzkopf et al., 2012). Our observations of correlated changes between the size of and cortical thickness in the primary visual cortices, and the microstructure of the visual pathways feeding into the cortex, may provide an extended mechanistic explanation of the structural basis for individual differences in visual processing and perception.

Alternatively, some of the multimodal components may be due to morphological differences such as folding patterns, which could have artifactual impact on GM measures and adjacent WM measures. However, all of the modalities attempt to correct for gross morphological differences in their own ways, which should be relatively independent from one another. In this case, adjacency would be the key factor and the direction of the association would be less certain. We emphasize that there is no constraint in the decomposition model

to ensure spatial correspondence across modalities. In this "flat linked ICA" configuration, the model is unaware of any spatial correspondence between modalities.

Reliability of the decomposition

We have shown that our approach yields components that are mainly robust to population subsampling and to changes in initialization. The global components are highly reliable, and most of the local components are also much more reliable than their PCA equivalents. It should be possible to more rigorously assess which components are *individually* significant, potentially using the null model derived in Groppe et al. (2009). As we aim to describe *classes* of components rather than attach interpretations to individual components, we leave this for future work.

General comments on Linked ICA

As this multimodal analysis revealed several distinct categories of components, e.g. global versus local, multimodal versus unimodal, and GM-dominated versus WM-dominated, they are not generally comparable to each other as modes of variation in the same way that resting state networks (Beckmann et al., 2005) or FA components (Li et al., 2011) are. The fact that many of the components are dominated by one modality suggests that it might be better to analyze the diffusion data separately from the morphometric data, and attempt to pair or cluster the components afterwards. However, the significant number of joint components suggests that there is value in a joint decomposition. Furthermore, we have shown that the current format is fully capable of isolating true single-modality components, such as artifactual differences in DTI. The decision to run a joint or separate analysis will ultimately depend on what questions are being asked of the data.

A theoretical benefit of linked ICA is the ability to perform automatic dimensionality selection on the fly, during the component inference. However in this paper we have used a restricted dimensionality; this was partially for computational reasons, but also because we believe that using a very high dimensionality can reduce interpretability by splitting components, even if this splitting is justified by the data (i.e. not overfitting to noise). Our previous work on FMRI has shown that there may be no single "correct" dimensionality for an ICA decomposition, but instead different dimensionalities will show features on a different scale (Smith et al., 2009). In this work we have chosen a dimensionality that we believe yields highly interpretable components of multimodal inter-subject variability; a higher dimensionality may give future studies more detail and allow them to answer more specific questions.

In each linked ICA component, all modalities share exactly the same subject loadings, and this is perhaps an overly-strong constraint. However, it is not clear how this linkage can be weakened without compromising interpretability. Regularization would be needed to encourage the subject loadings to be similar, and this must be carefully balanced against the pressure to make each spatial map as independent (and non-Gaussian) as possible. Furthermore, the resulting components would have no consensus subject loading vector, which introduces ambiguity into the post-hoc correlation analysis and leads to the open question of how to interpret the *differences* between these slightly different vectors. Our approach of using a hard linkage between modalities is conceptually and methodologically cleaner, at the expense of possibly requiring more components.

The variety in the types of components may be an indication that the current model is too flexible and allows too much variance to be absorbed by a small number of "global" components. This is potentially a different behavior from standard FastICA, which removes scale information (by prewhitening the data) before maximizing non-Gaussianity. A model based explicitly on sparsity (Li et al., 2010) could potentially provide a different set of components with less of a global/local divide, while maintaining the flexibility of a generative Bayesian framework. However, we leave any comparison between standard FastICA and Bayesian ICA methods for future work.

The current approach uses no information about the spatial relationship between the modalities. This is so that the model does not automatically induce spatially adjacent GM and WM patterns. However, potentially greater sensitivity could be obtained by also pooling spatial information, but there are obvious technical and conceptual difficulties in determining spatial correspondence between different modalities including e.g. cortical surface maps and the TBSS skeleton.

On a general note, despite efforts to recruit healthy participants only, including health interview, cognitive assessments and radiological evaluation, we cannot completely rule out the influence of subclinical conditions on the structural imaging measures. Further, we did not exclude areas showing T2-weighted white matter hyperintensities (WMHs). However, signal intensity alterations are regularly found in healthy individuals, and may not represent a specific neuroradiological marker of disease (Vernooij et al., 2007). Also, WMHs may manifest as characteristic patterns across the included modalities, and may therefore be isolated in specific components. For example, component #52 (Fig. 6d) reflects structural variability in a subset of the oldest participants and is strongly associated with increased MD and decreased GM density in areas including the caudate and putamen and associated WM pathways. Thus, we think that the present method alleviates some of the concerns related to inclusion of WMHs in the analysis.

Future avenues

A major future direction for development of this approach is to make Linked ICA work with a wider variety of modalities; e.g. surface based measures of gyrification and curvature, subcortical volumes, WM hyperintensities, signal intensity and contrast measures (Panizzon et al., 2012; Salat et al., 2009; Westlye et al., 2010a, 2009b). In particular, functional MRI could be incorporated in a number of ways, but must be summarized using regression parameter estimates for task-based data (Duff et al., 2012), or subject-specific RSN maps e.g. from dual regression (Filippini et al., 2009) for resting-state data. Longitudinal datasets also present a new direction, as there are many possibilities for modeling the baseline and change over multiple timepoints. Linked ICA may also provide an elegant way of generating endophenotypes for genetic association studies. Similarly, we plan to test whether these components are sensitive to disease states in multimodal datasets.

Of the major global components we found, several were strongly age-related. We believe this is likely because of the large age-span of our sample and the relative homogeneity of the group in most other respects. It would be interesting to test this on data sets with other large sources of variability (e.g. disease groups) to determine how much these components' spatial maps change as a result of a change in the population, and how well the subjects can be classified using the subject loading matrix as features. We also plan to investigate the question of dimensionality estimation more closely, keeping in mind that Bayesian model energy may not be ideal if the goal is to maximize reproducibility, interpretability, or relevance to behavioral/genetic variables.

In conclusion, Linked ICA provides an efficient and sensitive data-driven approach for defining components of inter-subject variability in neuroimaging data across various modalities. We propose that fusing variance across modalities may increase sensitivity to external subject variables and will focus on some aspects of the remaining components in other papers.

Acknowledgments

We are grateful to Anders Dale for the spline-fitting code, Stamatios Sotiropoulos for help in interpreting DTI components, and Christian

Beckmann for helpful conversations on ICA. We would also like to acknowledge funding from the Research Council of Norway to L.T.W. (204966), K.B.W. (177404/W50, 186092/V50, 204935/F20) and A.M.F. (189507/V40, 199537).

Appendix A. Contributions to shared subject-series

Our earlier paper (Groves et al., 2011) gave the variational Bayes full update equations needed to implement Linked ICA. Here we reproduce the updates for the shared subject-series matrix and show how it relates to the results shown in Fig. 5.

The VB approximate-posterior distribution $q()$ of the components-by-subjects matrix $\mathbf{H}$ is given by the multivariate normal distribution on each column

$$q(\mathbf{H}_r) = N(\mathbf{H}_r; \mathbf{M}_r, \mathbf{V}_r)$$

with the covariance and mean defined by

$$\mathbf{V}_{\mathrm{r}}^{-1} = \mathbf{I} + \Sigma_{k=1}^{K} \left\langle \lambda_{\mathrm{r}}^{(\mathrm{k})} \right\rangle \left\langle \mathbf{X}^{(\mathrm{k})T} \mathbf{X}^{(k)} \right\rangle \circ \left\langle \mathbf{W}^{(\mathrm{k})T} \mathbf{W}^{(\mathrm{k})} \right\rangle$$

$$\mathbf{M}_{\mathrm{r}} \mathbf{V}_{\mathrm{r}}^{-1} = \Sigma_{k=1}^{K} \left\langle \lambda_{\mathrm{r}}^{(\mathrm{k})} \right\rangle \mathbf{Y}_{\mathrm{r}}^{(\mathrm{k})T} \left\langle \mathbf{X}^{(\mathrm{k})} \right\rangle \circ \left\langle \mathbf{W}^{(\mathrm{k})} \right\rangle$$

where $k = 1 \ldots K$ indexes the modalities, $\mathrm{X}^{(\mathrm{k})}$ is a voxels-by-components matrix of spatial maps, $\mathrm{W}^{(\mathrm{k})}$ is a 1-by-components matrix that absorbs the overall component scaling, $\mathrm{Y}^{(\mathrm{k})}$ is the voxels-by-subjects data matrix, and $\boldsymbol{\lambda}^{(\mathrm{k})}$ is the noise precision of each subject. A subscript r indicates that only the column of the matrix corresponding to subject r is taken. Angle brackets indicate an expectation integrated over the VB posterior distribution of the contained random variables.

These summations mean that the inferred $\mathbf{H}$ is given by a precision-weighted combination of normal distributions from each of these K modalities (and unit prior), i.e.

$$N(\mathbf{H}_r; \mathbf{M}_r, \mathbf{V}_r) \propto N(\mathbf{H}_r; \mathbf{0}, \mathbf{I}) \; \Pi_{k=1}^{K} N\left(\mathbf{H}_r; \mathbf{M}\mathrm{part}_r^{(k)}, \mathbf{V}\mathrm{part}_r^{(k)}\right)$$

with the covariance parts and mean parts defined by:

$$\mathbf{V}\mathrm{part}_r^{(k)-1} = \left\langle \lambda_r^{(k)} \right\rangle \left\langle \mathbf{X}^{(k)T} \mathbf{X}^{(k)} \right\rangle \circ \left\langle \mathbf{W}^{(k)T} \mathbf{W}^{(k)} \right\rangle$$

$$\mathbf{M}\mathrm{part}_r^{(k)} \mathbf{V}\mathrm{part}_r^{(k)-1} = \left\langle \lambda_r^{(k)} \right\rangle \mathbf{Y}_r^{(k)T} \left\langle \mathbf{X}^{(k)} \right\rangle \circ \left\langle \mathbf{W}^{(k)} \right\rangle.$$

$\mathbf{M}\mathrm{part}(\mathrm{k})$ is gives modality k's contribution to the shared subject-course, and the diagonal elements of $(\mathbf{V}\mathrm{part}(\mathrm{k}))-1$ give the precision contributions of modality k.

References

Andersson, J.L.R., Jenkinson, M., Smith, S., 2007a. Non-linear optimisationFMRIB technical report TR07JA1 from www.fmrib.ox.ac.uk/analysis/techrep.

Andersson, J.L.R., Jenkinson, M., Smith, S., 2007b. Non-linear registration, aka spatial normalisation FMRIB technical report TR07JA2 from www.fmrib.ox.ac.uk/analysis/techrep.

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry—the methods. NeuroImage 11, 805–821.

Beck, A.T., Steer, R., 1987. Beck Depression Inventory Scoring Manual. The Psychological Corporation, New York.

Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans. Med. Imaging 23, 137–152.

Beckmann, C.F., DeLuca, M., Devlin, J.T., Smith, S.M., 2005. Investigations into resting-state connectivity using independent component analysis. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360, 1001–1013.

Buckner, R.L., Head, D., Parker, J., Fotenos, A.F., Marcus, D., Morris, J.C., Snyder, A.Z., 2004. A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. NeuroImage 23, 724–738.

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

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. NeuroImage 9, 179–194.

Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26, 55–67.

Dosenbach, N.U., Nardos, B., Cohen, A.L., Fair, D.A., Power, J.D., Church, J.A., Nelson, S.M., Wig, G.S., Vogel, A.C., Lessov-Schlaggar, C.N., Barnes, K.A., Dubis, J.W., Feczko, E., Coalson, R.S., Pruett Jr., J.R., Barch, D.M., Petersen, S.E., Schlaggar, B.L., 2010. Prediction of individual brain maturity using fMRI. Science 329, 1358–1361.

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, 2375–2386.

Douaud, G., Jbabdi, S., Behrens, T.E., Menke, R.A., Gass, A., Monsch, A.U., Rao, A., Whitcher, B., Kindlmann, G., Matthews, P.M., Smith, S., 2011. DTI measures in crossing-fibre areas: increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. NeuroImage 55, 880–890.

Duff, E.P., Trachtenberg, A.J., Mackay, C.E., Howard, M.A., Wilson, F., Smith, S.M., Woolrich, M.W., 2012. Task-driven ICA feature generation for accurate and interpretable prediction using fMRI. NeuroImage 60, 189–203.

Ennis, D.B., Kindlmann, G., 2006. Orthogonal tensor invariants and the analysis of diffusion tensor magnetic resonance images. Magn. Reson. Med. 55, 136–146.

Filippini, N., MacIntosh, B.J., Hough, M.G., Goodwin, G.M., Frisoni, G.B., Smith, S.M., Matthews, P.M., Beckmann, C.F., Mackay, C.E., 2009. Distinct patterns of brain activity in young carriers of the APOE-epsilon4 allele. Proc. Natl. Acad. Sci. U. S. A. 106, 7209–7214.

Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. U. S. A. 97, 11050–11055.

Fischl, B., Sereno, M.I., Dale, A.M., 1999a. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. NeuroImage 9, 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, 272–284.

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, 70–80.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341–355.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb. Cortex 14, 11–22.

Fischl, B., Rajendran, N., Busa, E., Augustinack, J., Hinds, O., Yeo, B.T., Mohlberg, H., Amunts, K., Zilles, K., 2008. Cortical folding patterns and predicting cytoarchitecture. Cereb. Cortex 18, 1973–1980.

Fjell, A.M., Westlye, L.T., Greve, D.N., Fischl, B., Benner, T., van der Kouwe, A.J., Salat, D., Bjørnerud, A., Due-Tønnessen, P., Walhovd, K.B., 2008. The relationship between diffusion tensor imaging and volumetry as measures of white matter properties. NeuroImage 42, 1654–1668.

Fjell, A.M., Walhovd, K.B., Westlye, L.T., Ostby, Y., Tamnes, C.K., Jernigan, T.L., Gamst, A., Dale, A.M., 2010. When does brain aging accelerate? Dangers of quadratic fits in cross-sectional studies. NeuroImage 50, 1376–1383.

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12, 189–198.

Franke, K., Ziegler, G., Kloppel, S., Gaser, C., 2010. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. NeuroImage 50, 883–892.

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, 21–36.

Groppe, D.M., Makeig, S., Kutas, M., 2009. Identifying reliable independent components via split-half comparisons. NeuroImage 45, 1199–1211.

Groves, A.R., Beckmann, C.F., Smith, S.M., Woolrich, M.W., 2011. Linked independent component analysis for multimodal data fusion. NeuroImage 54, 2198–2217.

Hasan, K.M., Sankar, A., Halphen, C., Kramer, L.A., Brandt, M.E., Juranek, J., Cirino, P.T., Fletcher, J.M., Papanicolaou, A.C., Ewing-Cobbs, L., 2007. Development and organization of the human brain tissue compartments across the lifespan using diffusion tensor imaging. Neuroreport 18, 1735–1739.

Hasan, K.M., Iftikhar, A., Kamali, A., Kramer, L.A., Ashtari, M., Cirino, P.T., Papanicolaou, A.C., Fletcher, J.M., Ewing-Cobbs, L., 2009. Development and aging of the healthy human brain uncinate fasciculus across the lifespan using diffusion tensor tractography. Brain Res. 18, 67–76.

Imperati, D., Colcombe, S., Kelly, C., Di Martino, A., Zhou, J., Castellanos, F.X., Milham, M.P., 2011. Differential development of human brain white matter tracts. PLoS One 6, e23437.

Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage 17, 825–841.

Johnstone, I.M., 2001. On the distribution of the largest eigenvalue in principal components analysis. Ann. Stat. 29, 137–152.

Lebel, C., Gee, M., Camicioli, R., Wieler, M., Martin, W., Beaulieu, C., 2012. Diffusion tensor imaging of white matter tract evolution over the lifespan. NeuroImage 60, 340–352.

Li, Y.-O., Mukherjee, P., Nagarajan, S., Attias, H., 2010. A novel variational Bayesian method for spatiotemporal decomposition of resting-state fMRI. ISMRM 18th Annual Meeting, Stockholm, Sweden, 1–7 May 2010. http://cds.ismrm.org/protected/10MProceedings/files/3500_7300.PDF.

Li, Y.O., Yang, F.G., Nguyen, C.T., Cooper, S.R., Lahue, S.C., Venugopal, S., Mukherjee, P., 2011. Independent component analysis of DTI reveals multivariate microstructural correlations of white matter in the human brain. Hum. Brain Mapp.

Lövdén, M., Laukka, E., Rieckmann, A., Kalpuzos, G., Li, T.-Q., Jonsson, T., Wahlund, L.-O., Bäckman, L., 2012. The Dimensionality of Between-Person Differences in White Matter Microstructure in Old age. Hum. Brain Mapp. http://dx.doi.org/10.1002/hbm.21518.

Østby, Y., Tamnes, C.K., Fjell, A.M., Westlye, L.T., Due-Tønnessen, P., Walhovd, K.B., 2009. Heterogeneity in subcortical brain development: a structural MRI study of brain maturation from 8–30 years. J. Neurosci. 29, 11772–11782.

Panizzon, M.S., Fennema-Notestine, C., Kubarych, T.S., Chen, C., Eyler, L.T., Fischl, B., Franz, C.E., Grant, M.D., Hamza, S., Jak, A., Jernigan, T.L., Lyons, M.J., Neale, M.C., Prom-Wormley, E.C., Seidman, L., Tsuang, M.T., Wu, H., Dale, A.M., Kremen, W.S., 2012. Genetic and environmental influences of white and gray matter signal contrast: A new phenotype for imaging genetics? NeuroImage 60 (3), 1686–1695.

Penke, L., Munoz Maniega, S., Murray, C., Gow, A.J., Hernandez, M.C., Clayden, J.D., Starr, J.M., Wardlaw, J.M., Bastin, M.E., Deary, I.J., 2010. A general factor of brain white matter integrity predicts information processing speed in healthy older people. J. Neurosci. 30, 7569–7574.

Reese, T.G., Heid, O., Weisskoff, R.M., Wedeen, V.J., 2003. Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. Magn. Reson. Med. 49 (1), 177–182.

Salat, D.H., Lee, S.Y., van der Kouwe, A.J., Greve, D.N., Fischl, B., Rosas, H.D., 2009. Age-associated alterations in cortical gray and white matter signal intensity and gray to white matter contrast. NeuroImage 48, 21–28.

Scholz, J., Klein, M.C., Behrens, T.E., Johansen-Berg, H., 2009. Training induces changes in white-matter architecture. Nat. Neurosci. 12, 1370–1371.

Schwarzkopf, D.S., Song, C., Rees, G., 2011. The surface area of human V1 predicts the subjective experience of object size. Nat. Neurosci. 14, 28–30.

Schwarzkopf, D.S., Robertson, D.J., Song, C., Barnes, G.R., Rees, G., 2012. The frequency of visually induced gamma-band oscillations depends on the size of early human visual cortex. J. Neurosci. 32, 1507–1512.

Seeley, W.W., Crawford, R.K., Zhou, J., Miller, B.L., Greicius, M.D., 2009. Neurodegenerative diseases target large-scale human brain networks. Neuron 62, 42–52.

Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23 (Suppl. 1), S208–S219.

Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. NeuroImage 31, 1487–1505.

Smith, S.M., Johansen-Berg, H., Jenkinson, M., Rueckert, D., Nichols, T.E., Miller, K.L., Robson, M.D., Jones, D.K., Klein, J.C., Bartsch, A.J., Behrens, T.E., 2007. Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. Nat. Protoc. 2, 499–503.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Proc. Natl. Acad. Sci. U. S. A. 106 (31), 13040–13045.

Tamnes, C.K., Ostby, Y., Fjell, A.M., Westlye, L.T., Due-Tonnessen, P., Walhovd, K.B., 2010. Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. Cereb. Cortex 20, 534–548.

Vernooij, M.W., Ikram, M.A., Tanghe, H.L., Vincent, A.J.P.E., Hofman, A., Krestin, G.P., Niessen, W.J., Breteler, M.M.B., van der Lugt, A., 2007. Incidental findings on brain MRI in the general population. N. Engl. J. Med. 357, 1821–1828.

Wahl, M., Li, Y.O., Ng, J., Lahue, S.C., Cooper, S.R., Sherr, E.H., Mukherjee, P., 2010. Microstructural correlations of white matter tracts in the human brain. NeuroImage 51, 531–541.

Walhovd, K.B., Westlye, L.T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Agartz, I., Salat, D.H., Greve, D.N., Fischl, B., Dale, A.M., Fjell, A.M., 2011. Consistent neuroanatomical age-related volume differences across multiple samples. Neurobiol. Aging 32, 916–932.

Wechsler, D., 1999. Wechsler Abbreviated Scale of Intelligence. The Psychological Corporation, San Antonio, TX.

Westlye, L.T., Walhovd, K.B., Bjørnerud, A., Due-Tønnessen, P., Fjell, A.M., 2009a. Error-related negativity is mediated by fractional anisotropy in the posterior cingulate gyrus — a study combining diffusion tensor imaging and electrophysiology in healthy adults. Cereb. Cortex 19, 293–304.

Westlye, L.T., Walhovd, K.B., Dale, A.M., Espeseth, T., Reinvang, I., Raz, N., Agartz, I., Greve, D.N., Fischl, B., Fjell, A.M., 2009b. Increased sensitivity to effects of normal aging and Alzheimer's disease on cortical thickness by adjustment for local variability in gray/white contrast: a multi-sample MRI study. NeuroImage 47, 1545–1557.

Westlye, L.T., Walhovd, K.B., Dale, A.M., Bjornerud, A., Due-Tonnessen, P., Engvig, A., Grydeland, H., Tamnes, C.K., Ostby, Y., Fjell, A.M., 2010a. Differentiating maturational and aging-related changes of the cerebral cortex by use of thickness and signal intensity. NeuroImage 52, 172–185.

Westlye, L.T., Walhovd, K.B., Dale, A.M., Bjornerud, A., Due-Tonnessen, P., Engvig, A., Grydeland, H., Tamnes, C.K., Ostby, Y., Fjell, A.M., 2010b. Life-span changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry. Cereb. Cortex 20, 2055–2068.

Woolrich, M.W., Behrens, T.E., Beckmann, C.F., Jenkinson, M., Smith, S.M., 2004. Multilevel linear modelling for FMRI group analysis using Bayesian inference. NeuroImage 21, 1732–1747.

Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S.M., 2009. Bayesian analysis of neuroimaging data in FSL. NeuroImage 45, S173–S186.

Worsley, K.J., Poline, J.B., Vandal, A.C., Friston, K.J., 1995. Tests for distributed, nonfocal brain activations. NeuroImage 2, 183–194.

Zatorre, R.J., Fields, R.D., Johansen-Berg, H., 2012. Plasticity in gray and white: neuroimaging changes in brain structure during learning. Nat. Neurosci. 15, 528–536.

Ziegler, G., Dahnke, R., Jancke, L., Yotter, R.A., May, A., Gaser, C., 2011. Brain structural trajectories over the adult lifespan. Hum. Brain Mapp. http://dx.doi.org/10.1002/hbm.21374 (Electronic publication ahead of print 6 Sep 2011).