Normative development of white matter tracts: Similarities and differences in relation to age, gender and intelligence

Jonathan D. Clayden[1†], Sebastian Jentschke[1], Mónica Muñoz[1,2], Janine M. Cooper[1], Martin J. Chadwick[1,3], Tina Banks[4], Chris A. Clark[1] & Faraneh Vargha-Khadem[1]

[1]Institute of Child Health, University College London, 30 Guilford Street, London WC1N 1EH, UK; [2]School of Medicine, University of Castilla–La Mancha, Avenida Almansa 14, 02006 Albacete, Spain; [3]Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, 12 Queen Square, London WC1N 3BG, UK; [4]Great Ormond Street Hospital for Children, Great Ormond Street, London WC1N 3JH, UK

[†]*Corresponding author. E-mail: j.clayden@ucl.ac.uk*

The white matter of the brain undergoes a range of structural changes throughout development; from conception to birth, in infancy, and onwards through childhood and adolescence. Several studies have used diffusion magnetic resonance imaging (dMRI) to investigate these changes, but a consensus has not yet emerged on which white matter tracts undergo changes in the later stages of development, or what the most important driving factors are behind these changes. In this study of typically developing 8 to 16 year-old children, we use a comprehensive data-driven approach based on principal components analysis to identify effects of age, gender and brain volume on dMRI parameters, as well as their relative importance. We also show that secondary components of these parameters predict full-scale IQ, independently of the age and gender-related effects. This overarching assessment of the common factors and gender differences in normal white matter tract development will help to advance understanding of this process in late childhood and adolescence.

Introduction

Late childhood and adolescence are important phases in the normal development of the brain. While the human brain reaches 90% of its adult size by the age of five, structural change to both the grey and white matter—and hence the neural information processing circuits which they subserve—continues right up to the onset of adulthood (Tau & Peterson, 2010). In particular, myelination of neuronal axons is reported to be one key process which continues beyond childhood (Benes et al., 1994; Yakovlev & Lecours, 1967). However, it is not yet fully clear which white matter tracts undergo changes in this phase of development, or what the most important driving factors are behind these changes.

A number of studies have used diffusion magnetic resonance imaging (dMRI) to investigate the structural changes in neural white matter during childhood and adolescence. This technique provides microstructural biomarkers including fractional anisotropy (FA), broadly interpreted as a marker of the structural "integrity" of white matter; and mean diffusivity (MD), a directionally averaged measure of the degree of water self-diffusion in tissue (Basser & Pierpaoli, 1996). In general, FA tends to increase with age while MD decreases, most likely reflecting the ongoing process of myelination, but synaptogenesis and synaptic pruning also continue well beyond infancy and may play a role. However, specific observations vary due to differing study populations and methodological details. White matter structures variously reported to show changes during childhood and adolescence include the arcuate fasciculi, corpus callosum, corticospinal tracts, inferior longitudinal fasciculi and uncinate fasciculi (Ashtari et al., 2007; Barnea-Goraly et al., 2005; Bonekamp et al., 2007; Eluvathingal et al., 2007; Giorgio et al., 2010; Lebel et al., 2008; Muetzel et al., 2008; Schmithorst et al., 2002), although most studies have used whole-brain voxel-based analysis methods to observe these effects. Standard structural imaging has also been used to demonstrate changes in white matter density with age (Paus et al., 1999), and cognitive development has been shown to be linked to the microstructural changes identified with dMRI (Mabbott et al., 2006; Nagy et al., 2004).

There has, however, been little evidence presented of substantive differences between the genders, despite

findings of sexual dimorphism in the white matter of human adults and of rodents (Cerghet et al., 2006; Gur et al., 1999; Kim & Juraska, 1997; Leonard et al., 2008). The most consistently reported morphological difference between the sexes in humans is of approximately 8–12% greater total brain volume in males (Lenroot & Giedd, 2010), but there is also evidence that grey and white matter volumes vary differently with age between the genders (De Bellis et al., 2001; Blanton et al., 2004; Caviness et al., 1996). Schmithorst et al. (2008) presented an analysis of dMRI data from a group of 5–18 year olds, indicating some age–gender interactions in FA and MD, although their analysis was performed over a single large white matter mask, with the magnitude of the effect in specific brain regions established using post-hoc statistical tests. More recently, Asato et al. (2010) found some age group–gender interactions in radial diffusivity, another measure derived from dMRI, using three age groups representing children, adolescents and young adults. Since these studies looked at different diffusion measures and represented age differently, replication of age–gender interaction effects is still lacking, and the extent of gender differences in white matter development trajectories remains an open question.

Recent work has also demonstrated that there is a substantial link between dMRI-based microstructural parameters measured from different tracts, going well beyond an interhemispheric correspondence in bilateral structures (Wahl et al., 2010). Westlye et al. (2010) have demonstrated substantial correlations between tracts across the life span; and a common factor of white matter integrity—based on measurements from a set of eight major tracts—has been shown to predict information processing speed across a group of healthy older people of very similar age (Penke et al., 2010). These studies strongly suggest that there is a significant global component to the variance in these parameters, which collectively has clear functional relevance.

Evidence for a link between dMRI-based biomarkers and general intelligence, in some localised regions of white matter, was presented by Schmithorst et al. (2005). In addition, Tamnes et al. (2010) showed somewhat different patterns of relationship between verbal and performance ability measures, and microstructure parameters. Both of these previous studies used voxelwise analysis methods which make no allowance for the commonality between tracts; but nevertheless, differences in intelligence between individuals are significant because of their association with variation in life outcomes, including occupational attainment and health (Deary et al., 2010). Neuroimaging is important in offering the opportunity to investigate their structural predictors.

In this work we take a tract-based approach to analysing microstructural development, using a recently developed, automated method to segment the white matter structures of interest (Clayden et al., 2009b). Unlike most previous studies, we do not isolate one or two regressors of interest, but rather use data-driven principal components analysis to identify common factors, and then link these components to age, gender, brain volume and intelligence measures in a robust and structured prospective analysis. This approach allows us to identify groups of tracts which "vary together" and analyse them as a set, thereby dramatically improving data efficiency relative to voxel-by-voxel or tract-by-tract approaches. We find substantial evidence for differences in the rates of change of dMRI parameters between the sexes. Variance in FA and MD across the group is found to be driven primarily by age-related changes, but the total volume of the brain is also found to relate to these measures. Moreover, factors in FA and MD which are secondary to the age effect independently predict intelligence. Taken together, this comprehensive analysis helps to advance understanding of the key factors in white matter development during late childhood and adolescence.

Methods

The participants for this study were 59 healthy children (34 female), with ages ranging from 8 to 16 yr (mean 11.5 yr, standard deviation 2.1 yr; boys' mean 10.9 yr, girls' mean 11.9 yr). These children were recruited by advertising in local schools, or through relationships to children being scanned as part of a clinical study. A parent of each participant answered a standardised questionnaire, the Child Behaviour Checklist (http://www.aseba.org), to identify and exclude children with a potential behavioural or psychiatric problem. Any contraindication to MRI was also grounds for exclusion. The study was approved by the local ethics committee, and informed consent was obtained from both participants and parents.

Each participant underwent a dMRI protocol on a Siemens Avanto 1.5 T clinical system (Siemens Healthcare, Erlangen, Germany), using a self-shielding gradient set with maximum gradient strength of 40 mT m^{-1}, and standard "birdcage" quadrature head coil. Echo-planar diffusion weighted images were acquired for an isotropic set of 20 noncollinear directions, using a weighting factor of $b = 1000$ s mm^{-2}, along with a T_2-weighted ($b = 0$) volume. This protocol was repeated three times in a single scan session, and the data merged together without averaging. 45 contiguous axial slices of width 2.5 mm were imaged, using a field of view of 240×240 mm and 96×96 voxel acquisition matrix, for a final image resolution of $2.5 \times 2.5 \times 2.5$ mm. Echo time was 89 ms and repetition time was 6300 ms. In addition, a T_1-weighted 3D FLASH structural image was acquired using 176 contiguous sagittal slices, a 256×224 mm field of view and $1 \times 1 \times 1$ mm image resolution. Echo time in this case was 4.9 ms, and repetition time was 11 ms. Overall scan time for these sequences was approximately 19 minutes.

All 59 T_1-weighted structural images were used for the brain and white matter volume analysis. However, six dMRI data sets were removed due to significant motion artefacts in the diffusion-weighted images, leaving a set of 53 participants (30 female) for tractography.

In addition to brain imaging, full-scale IQ values were obtained for each participant using the Wechsler Intelligence Scale for Children, fourth UK edition (WISC-IV UK).

Image processing

DICOM format image files obtained from the scanner were converted to NIfTI-1 format using the TractoR software package (http://www.tractor-mri.org.uk; Clayden et al., 2011). Diffusion data were preprocessed to correct for eddy-current induced distortions using tools in the FSL package (http://www.fmrib.ox.ac.uk/fsl). The brain was segmented using FSL's brain extraction tool (Smith, 2002), and diffusion tensors were calculated using least-squares fitting.

Probabilistic neighbourhood tractography (PNT), as implemented in TractoR, was applied to segment the arcuate fasciculi, anterior thalamic radiations (ATRs), cingulum bundles, corticospinal tracts (CSTs), inferior longitudinal fasciculi (ILFs), uncinate fasciculi, and corpus callosum genu and splenium (forceps minor and major) in each participant. This method uses a statistical model of tract shape, and a set of reference tracts based on a published white matter atlas, to automatically and robustly segment tracts of interest in the native image space of each participant (Clayden et al., 2009b; Muñoz Maniega et al., 2008). It also provides a natural index of the "goodness" of the segmentation in each participant, which is independent of all diffusion parameters. FSL BEDPOSTX/ProbTrack, with up to two fibre directions modelled per voxel, was used as the underlying tractography algorithm (Behrens et al., 2007). False positive streamlines were removed using an automated pruning algorithm, based on the same tract shape modelling approach, as described previously (Clayden et al., 2009a). Mean FA and MD within the final segmentations, weighted by visitation count after pruning, were then calculated. No threshold on visitation count was applied.

The brain was also segmented from the T_1-weighted structural images using FSL's brain extraction tool. The brain extraction was repeated iteratively to improve robustness, and the total brain volume was calculated from the final segmentation (Fagiolo et al., 2008). FSL FAST was then applied to perform gross tissue classification into grey matter, white matter and cerebrospinal fluid (Zhang et al., 2001). A conservative threshold of 99% classification confidence was applied and the total volume of segmented white matter was recovered.

Statistical analysis

Due to occasional failures of the tractography algorithm, not every tract could be satisfactorily segmented in every participant. Failures were detected by performing a standard outlier detection on the "goodness-of-fit" measure provided by PNT. Specifically, data for a participant were ignored for a particular tract if the goodness-of-fit score was more than 1.5 times the interquartile range below the first quartile for that tract. No more than five data sets were ignored for any tract in this way.

To identify relationships between FA and MD measurements across tracts, principal components analysis (PCA) was performed on each measure separately using the Nonlinear Iterative PArtial Least Squares (NIPALS) approach (Wold, 1966), as implemented in the "pcaMethods" package for R, version 1.30.0 (Stacklies et al., 2007). We used this iterative technique because it can handle the missing values excluded due to poor segmentation, and therefore makes full use of the available data. PCA transforms the data into a set of "loadings", or levels of contribution from each tract, and "scores" for each principal component in each subject (Bratchell, 1989). A small number of major principal components can often be used as a set of summary variables which capture most of the variability in the data set. Scores for major principal components were further analysed for relationships to age, gender and total brain volume using ANCOVA. The value of the FA and MD principal component scores as linear predictors of full-scale IQ was investigated using ANCOVA and additional F-tests.

Total brain and white matter volumes were separately tested for age and gender effects using ANCOVA.

All statistical analysis was performed with R (R Development Core Team, 2010). Graphical representations of our results were created using the "ggplot2" R package, version 0.8.7 (Wickham, 2009).

Results

Group overlay maps of the tracts of interest are shown in Fig. 1. We note that tract segmentation consistency is high between participants using probabilistic neighbourhood tractography, with very few false positive pathways. There was no need to apply an arbitrary threshold to the streamline visitation maps.

The results of our tract and tissue volume analyses are laid out below.

Brain and white matter volumes

The mean (± standard deviation) total brain volume across the data set was 1465 ± 89 ml for girls, and 1582 ± 145 ml for boys. The average brain volume is therefore 8% higher in boys. Mean white matter volume was 404 ± 35 ml for girls and 437 ± 53 ml for boys, or

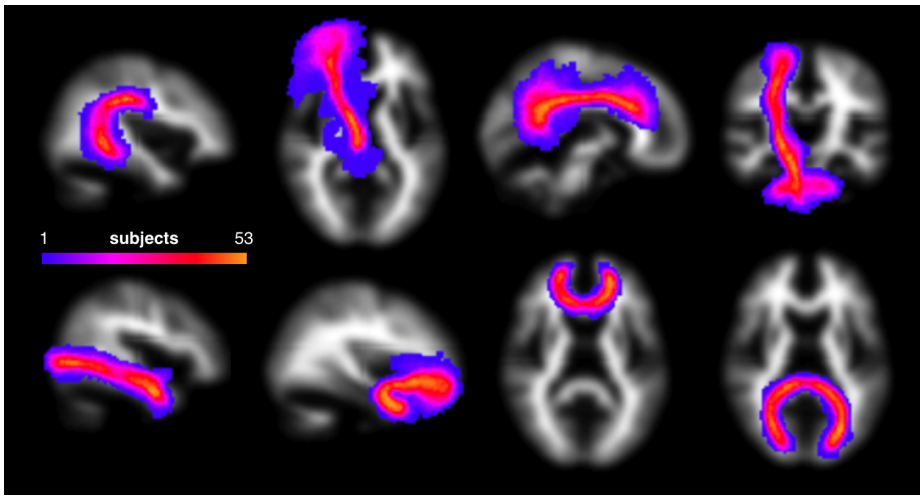


Figure 1: Group maps of the right-sided tracts of interest, plus forceps minor and major, after segmentation using PNT. The colour scale indicates the proportion of subjects whose segmentations overlap with each voxel in MNI standard space.

27.5 ± 1.3% and 27.6 ± 1.1% of total brain volume respectively.

Accordingly, we found that total brain volume in our data set was strongly dependent on gender ($F_{1,53}$ = 14.55, $P < 0.001$). Linear and quadratic main effect terms in age were also statistically significant, as were their interactions with gender (all $P < 0.05$).

While white matter volume was very strongly predicted by total brain volume ($F_{1,51}$ = 391.48, $P \ll 0.001$), there was also a significant main effect of age on white matter volume after brain volume was taken into account ($F_{1,51}$ = 8.83, $P < 0.01$), as well as an age–gender interaction ($F_{1,51}$ = 5.73, $P < 0.05$) and an age–brain volume interaction ($F_{1,51}$ = 6.10, $P < 0.05$). There was no evidence for a nonlinear age effect in this case, so these terms were not included in the statistical model.

The latter age–gender interaction indicates that white matter volume increases more slowly with age in boys than in girls, once the general effect of whole brain growth is taken into account. For illustration, the relationship between age and the proportion of white matter in the brain is shown for each gender in Supplementary Fig. 1, but given the very strong correlation between brain and white matter volumes, only a small amount of the variance in white matter volume is unexplained by brain volume, and so this proportion is expected to be rather noisy.

Tract diffusion characteristics

Scatter plots of FA and MD for each tract are shown in Figs 2 and 3, respectively, along with least-squares regression lines for each gender. Several major characteristics of the data are immediately obvious from these plots. There is clearly some difference in the diffusion parameters between tracts. On the whole, FA tends to increase or stay level with age, while MD decreases or stays level. The greatest variance in both FA and MD values is found in the corpus callosum splenium—in the case of MD, so much so that it was necessary to plot the data on their own scale. For MD, there is a consistent tendency for the trend line for boys to begin above that for girls at young ages, and then later to cross it. This suggests that changes in MD with age are gender-specific.

The evidence for different rates of change of the diffusion parameters between the two genders is highlighted in Fig. 4, which plots the estimated slopes for each tract–hemisphere–gender combination, along with their 95% confidence intervals. Notice that, to within these confidence bounds, the slopes fitted to data from girls do not differ from zero, with the marginal exceptions of the left arcuate and left ILF for FA. On the other hand, FA slopes are greater than zero for boys in bilateral ATRs, right cingulum and left CST; and MD slopes are less than zero in bilateral arcuate fasciculi, left ATR, bilateral cingula, bilateral CSTs, and marginally for right ILF and bilateral uncinate fasciculi. Tract-by-tract analysis using *t*-tests shows a significant difference between the sexes in left CST FA ($t_{45.9}$ = 2.80, $P < 0.01$) and MD ($t_{42.3}$ = −2.82, $P < 0.01$); and in right ILF MD ($t_{42.5}$ = −2.33, $P < 0.05$). However, it is clear from the raw data in Figs 2 and 3, as well as the fit information in Fig. 4, that there is a far more consistent trend in the slopes than these isolated significant values suggest. We will therefore focus below on using our principal components instead, thereby consolidating common information across tracts. Fits including quadratic terms in age were investigated, but there was no support for curvature in the data.

Principal components analysis

The scree plot in Fig. 5 shows the results of our PCA. For MD there is a clearly dominant first principal component (PC) which encapsulates 44% of the variance, indicating a significant degree of shared variability across the set of tracts of interest. In the case of FA the picture is less clear. The first PC covers just 23% of the variance, and

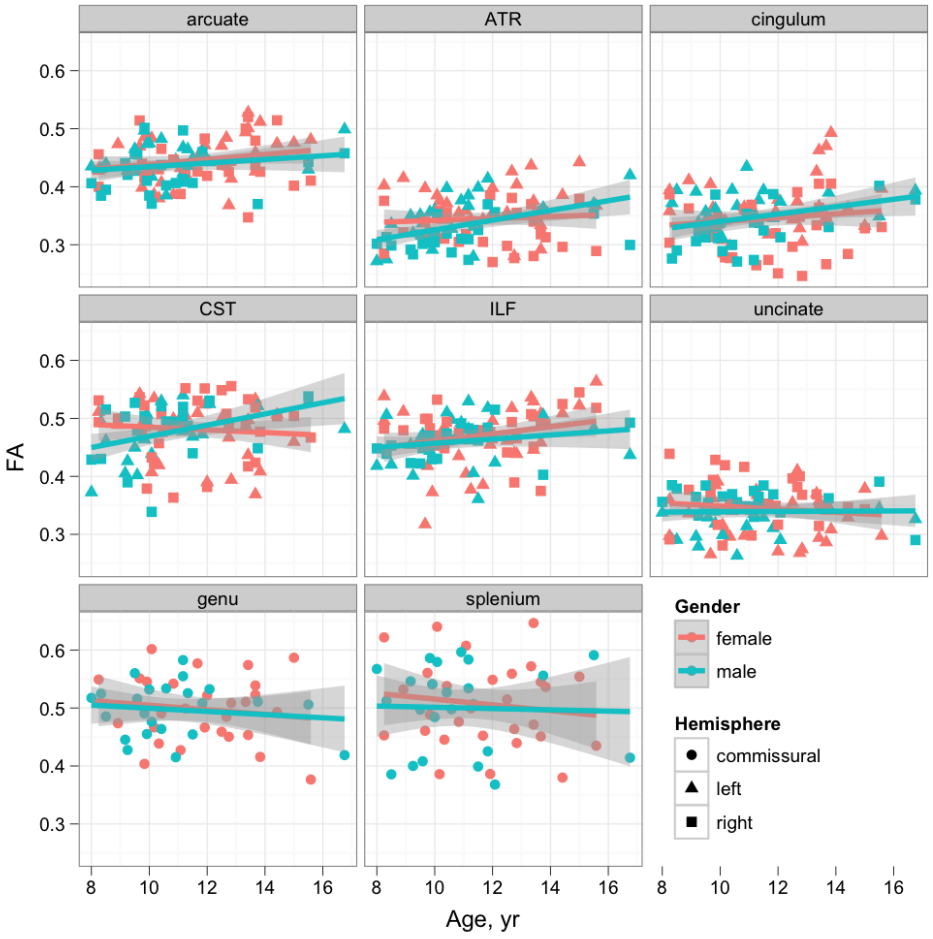


Figure 2: Scatter plots of age against FA for all tracts of interest. Linear regression lines and associated standard errors are shown for each gender.

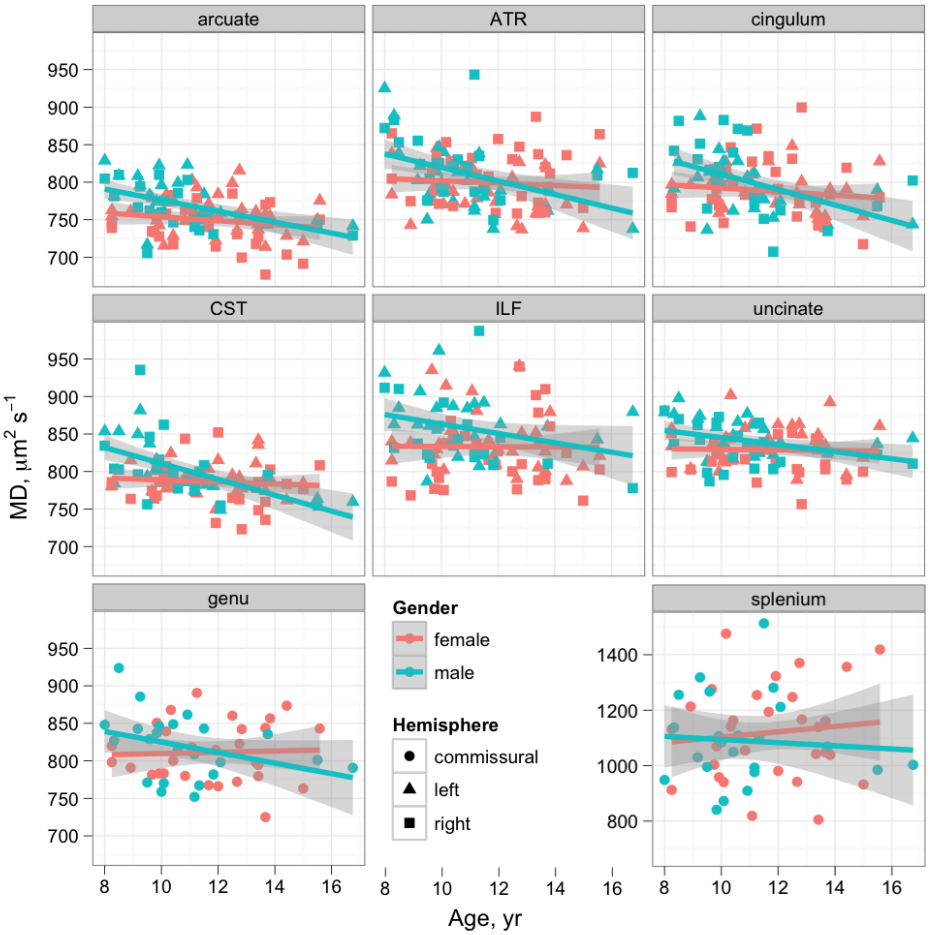


Figure 3: Scatter plots of age against MD for all tracts of interest. Linear regression lines and associated standard errors are shown for each gender. The splenium subplot uses a different y-axis to the others due to its much greater variability across individuals.

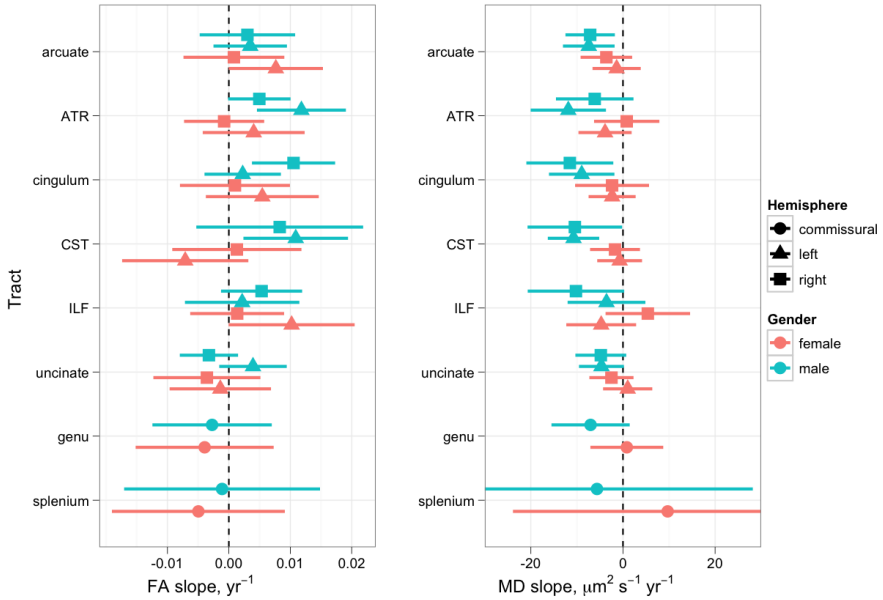


Figure 4: Estimated linear slope coefficients and their 95% confidence intervals for each tract–hemisphere–gender combination.

there are three components with more than 10%. These results reflect the lesser degree of similarity in data patterns seen across tracts in FA, as opposed to MD, in Figs 2 and 3. Jointly, components with more than 10% of the variance amount to 56% in MD and 48% in FA.

The contributions of the 14 tracts to each of the major principal components, or loadings, are shown in Fig. 6. We observe that all tracts show positive loadings on the first PC in both FA and MD—with the negligible exception of the genu in FA, whose loading is just barely negative. The second PC in FA is positively loaded for the arcuate fasciculi, cingula and left ILF, but negatively for all other tracts. The third PC in FA and the second in MD are loaded similarly to one another, and very strongly for the corpus callosum splenium.

Table 1 shows the results of our ANCOVA, based on the PCA scores for all components with more than 10% of the variance. F-values and P-values for age, gender and brain volume main effects, and age interactions, are shown for each component. Since white matter volume is very strongly correlated with total brain volume we did not include it as a separate term in the model. We observe that both FA and MD show significant main effects of age and brain volume, as well as a significant age–gender interaction, but for FA these effects are split across the two biggest components, whereas for MD they all appear in the first. Gender–brain volume and three-way interaction terms were included in each model, but in no case were these effects significant. Scores are plotted against age in Supplementary Fig. 2, and differences in slope between the genders may be directly observed in FA PC2 and MD PC1. We investigated the smaller components in each case, but found no evidence that they would add anything to our understanding or interpretation of the data.

Prediction of IQ

The mean (± standard deviation) full-scale IQ amongst participants for whom all MRI data were available was 113 ± 11 (range 88–137).

ANCOVA including terms for all three major components in FA showed that only the third component predicted full-scale IQ ($F_{1,49} = 8.36$, $P < 0.01$). Likewise, the second principal component in MD predicted IQ ($F_{1,50} = 4.60$, $P < 0.05$), whilst the first did not.

To demonstrate that the major components capture the important variability in FA and MD with regard to predicting IQ, we compared the models containing just these components to ones including all principal components. No significant improvement in the explained variance was observed for either FA ($F_{11,38} = 0.33$, $P = 0.97$) or MD ($F_{12,38} = 1.20$, $P = 0.32$).

Discussion

In this study we have used a thorough, data-driven analysis approach to make the following major contributions. Firstly, we have confirmed previous reports of general FA increases and MD decreases with age over the range of our data set (Figs 2 and 3). Secondly, we have observed—almost invariably across the 14 tracts studied—faster rates of decrease in MD in boys relative to girls (Figs 4 and 6; Supplementary Fig. 2; Table 1). It appears likely that the rates of change for girls are essentially zero for most or all of the tracts we have examined. There is some evidence for a comparable age–gender interaction in FA, but it is more tract-specific (Fig. 6; Table 1). Thirdly, total brain volume has been shown to depend on gender and (nonlinearly) on age. Total white matter volume depends on brain volume, unsurprisingly, but it

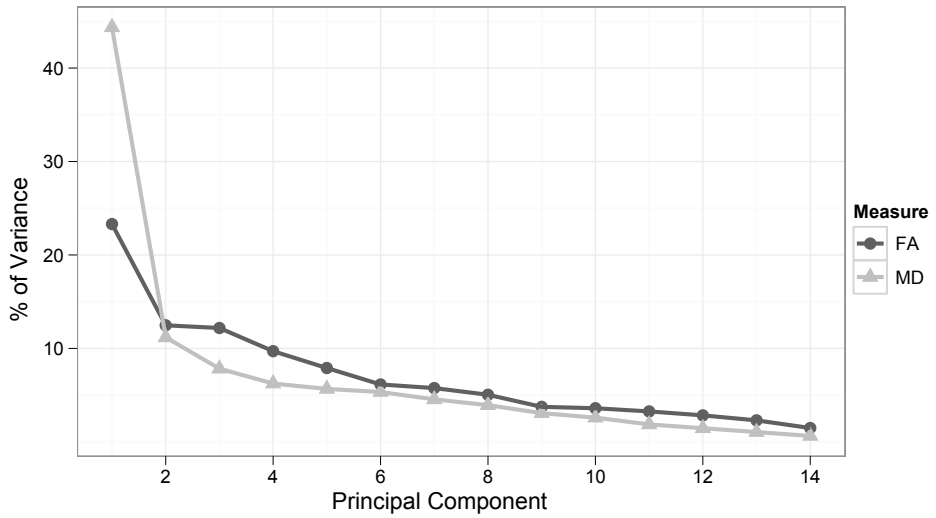


Figure 5: Scree plot showing the proportion of the variance attributed to each factor in our PCA, performed separately for FA and MD.

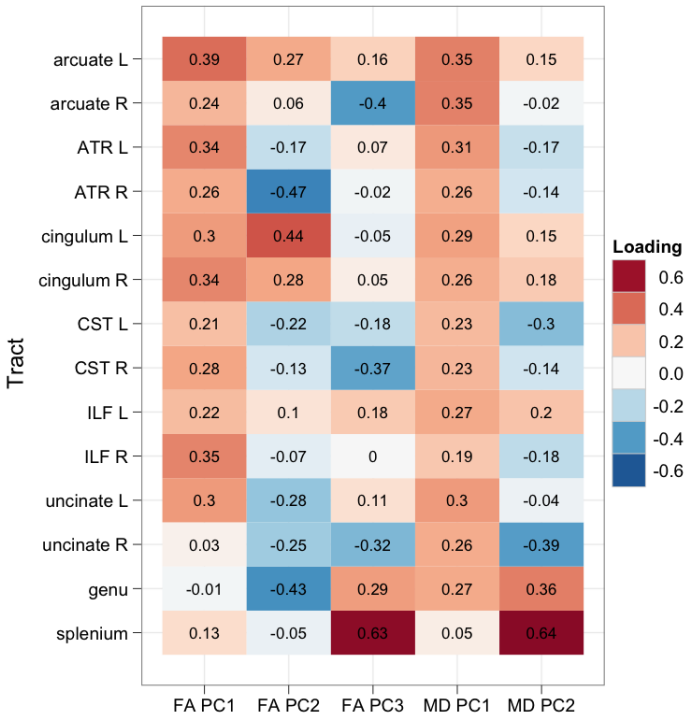


Figure 6: Colour-coded loading table for each principal component (PC) with greater than 10% of the variance. Greater loadings indicate greater contribution of the tract to the relevant PC.

	Age		Gender		TBV		Age × Gender		Age × TBV	
	$F_{1,45}$	P	$F_{1,45}$	P	$F_{1,45}$	P	$F_{1,45}$	P	$F_{1,45}$	P
FA PC1	**7.291**	**0.010**	1.122	0.295	0.103	0.749	0.012	0.915	0.652	0.423
FA PC2	0.706	0.405	0.017	0.898	**5.006**	**0.030**	**5.050**	**0.030**	2.918	0.095
FA PC3	0.012	0.915	0.147	0.703	0.117	0.734	0.146	0.704	<0.001	0.996
MD PC1	**17.245**	**<0.001**	0.060	0.807	**17.557**	**<0.001**	**5.666**	**0.022**	1.351	0.251
MD PC2	0.392	0.535	0.245	0.623	0.026	0.872	0.303	0.585	0.124	0.726

Table 1: Major ANCOVA results for each principal component (PC) with greater than 10% of the variance. F-values and P-values are given for the effects of age, gender and total brain volume (TBV), along with first-level interactions involving age. Significant values ($P < 0.05$) are shown in bold.

also changes with age, with some difference in the trajectory between the sexes. Fourthly, using principal components analysis we have reported that a significant amount of variance in FA and (particularly) MD is shared between tracts, and the largest principal component in both cases correlates with age (Fig. 5; Table 1). There are also significant effects of brain volume on FA and MD principal components, perhaps reflecting a different packing of axons and myelin in larger brains. Finally, we have demonstrated that principal components in FA and MD without links to age or gender independently predict intelligence. Significantly, all of these observations have been made using a structured, prospective analysis on a single data set, by identifying a set of major factors in each parameter, thereby taking shared variance into account and maximising data efficiency.

The age–gender interaction for MD was observed to be highly consistent across tracts. By inspection of Fig. 3, it seems that the gender difference in slope may be driven primarily by the younger end of the age group, where boys generally have higher MD than girls. Indeed, we found that restricting the analysis to just the 41 subjects (21 female) with age less than 13 yr had very little effect on the pattern of estimated slopes in Fig. 4, apart from broadening the confidence intervals slightly as would be expected with fewer observations (data not shown). It is plausible that the two slopes would be much more similar beyond the upper limit of our age range, but we found no support for curvature in the trajectories within the range studied here. Likewise, it might be hypothesised that the decline in MD observed specifically in boys would be observed in girls too, but at a younger age, reflecting the tendency for earlier development in females. However, fully testing this hypothesis would require significantly more data to be acquired from younger children and is therefore left to future work. For FA the picture is more mixed (see Fig. 2). In general, boys have a more positive rate of change with age, but the overall effect is less coherent—a view which was borne out by the PCA, as discussed below.

The finding of firm age–gender interactions in dMRI parameters is not unprecedented, as Schmithorst et al. (2008) have previously reported similar findings. Directly comparing the results of the two studies is not easy due to their very different image analysis and statistical methodologies, as well as the broader age range of the earlier study, but the message of difference in the rates of change of microstructural development between the sexes is common to both. Nevertheless, our approach has the advantage of using only planned statistics based on a structured linear model, rather than relying on post-hoc analysis to determine the magnitudes of the effects. We have also employed a well-tested and objective method for segmenting tracts of interest, whose reproducibility has been previously demonstrated (Clayden et al., 2009b).

Tanner staging was not performed for this study: instead, we have concentrated on relating imaging metrics to chronological age, since many developmental landmarks are still very much defined in these terms. (Age also has the advantages of being absolute and quantitative.) Nevertheless, linking pubertal status to these metrics would be an interesting approach for future work, and would shed light on the extent to which the white matter development trajectories observed depend on the relative onset ages for adolescence in the two sexes.

Although some previous studies have described lateralisation effects in developing white matter (e.g. Bonekamp et al., 2007; Eluvathingal et al., 2007), we found no strong evidence for interhemispheric differences in the development process. Fig. 4 shows some minor differences between left and right slopes, but the effect is much less marked than the contrast between the sexes.

Total brain volume was found to be an average of 8% higher in boys than girls, in agreement with previous findings (Lenroot & Giedd, 2010). It was found to depend nonlinearly on age, with some gender difference in trajectory, but the changes over this age range are fairly slight since the brain is very close to its adult size at the lower extreme of the range, with the early phase of rapid growth being already complete (Groeschel et al., 2010). Total white matter volume was observed to increase very slightly with age, and more rapidly in girls than boys, after taking total brain volume into account (see Supplementary Fig. 1).

Our principal components analysis is a major novel component of this study, and provides advantages in terms of data consolidation, multiple testing and the isolation of "layered" independent effects. The effect of this approach is to identify groups of tracts which "vary together" across subjects, calculate a weighted average FA or MD value for each group—where the loadings represent the weights—and perform subsequent statistical analysis on these averages. This approach is efficient with data and reduces measurement error by averaging, while still allowing us, through the loadings, to identify which gender and age effects apply to which tracts. Whilst the combination of PCA loadings and factor scores is a slightly unconventional way of presenting this kind of data, it is a robust approach whose results are extremely informative if correctly interpreted (Bratchell, 1989). PCA has been used for discrimination analysis and data dimensionality reduction in a few previous diffusion MRI studies (Caprihan et al., 2008; Lin et al., 2006; Robinson et al., 2010; Teipel et al., 2007), but as far as we are aware, only one previous study has used this powerful method for studying FA and MD data from a set of white matter tracts (Penke et al., 2010).

The largest principal component, or factor, which encapsulates 44% of the total variance in MD and 23% in FA, is linked to age (see Supplementary Fig. 2 and Table 1). The loadings show that this factor relates to all tracts except splenium for MD; and all except right uncinate, genu and splenium for FA (Fig. 6). For MD, this component also showed a significant relationship to total brain volume, and a significant age–gender interaction effect. Since this component is loaded for almost every tract, we conclude that the difference in MD slope between the genders is essentially ubiquitous across the tracts we have studied. Supplementary Fig. 2 confirms that the slope is more negative for males than for females. The effect of brain volume—which cannot be simply due to the difference between genders, because gender appeared as a separate term in our model—may reflect different packing of axons and myelin in different sized brains.

The second PC in FA, rather than the first, showed the age–gender interaction that indicates differing maturation trajectories between the sexes. The loadings on this factor were less consistent, being positive only for the association fibres of the cingulum bundles, arcuate fasciculi and left ILF. Note, however, that these are exactly the tracts where a *less* positive slope was observed for boys relative to girls (see Fig. 4); with all other tracts showing the opposite effect, and having a correspondingly negative loading on this factor. It is therefore clear that the age–gender interaction is less consistent across the brain for FA than it is for MD. This may be partly because FA is intrinsically noisier, being based on differences rather than a simple mean (Basser & Pierpaoli, 1996), but likely also reflects different development paths for these association tracts.

The third PC in FA and second PC in MD independently predict general intelligence, as measured by full-scale IQ derived from the WISC-IV UK. Given their analogy in this sense, it is not surprising that their loading patterns are similar, with both having heavy positive loading on the corpus callosum, particularly the splenium. Notable positive loadings are also observed in the left-sided inferior longitudinal and arcuate fasciculi. This is a powerful finding, suggesting that microstructural integrity has a role in between-individual variability in childhood intelligence. Since these components are not linked to age or gender, the impact of white matter integrity on IQ seems to be independent of the broader influence of those factors. Furthermore, other work has suggested that childhood IQ is linked with both IQ and white matter integrity in old age (Deary et al., 2006), so these effects may have significant long-term consequences. We note that while one previous study looking at microstructural predictors of IQ found little evidence of a link with MD (Schmithorst et al., 2005), statistical power and multiple comparison correction issues may have contributed to that negative finding. The mean IQ in our data set is slightly above the population average, but the range nevertheless covers more than three times the population standard deviation of 15. A recent adult twin study has demonstrated that genetic factors can mediate links between white matter integrity and intelligence (Chiang et al., 2009), and investigation of genetic influences in childhood would be a very constructive avenue for further work.

In this context, where tract characteristics are highly correlated, PCA is an extremely data-efficient approach to the analysis. A retrospective power calculation indicates that in the ATRs, for example—which show a clear difference between genders in Figs 2 and 4, and are substantially loaded on FA PC2—approximately 101 subjects (for the left ATR) or 107 (for the right) would be required *of each gender* in order to find a significant effect at $P < 0.05$ with 80% power in an individual tract analysis, assuming that our results are representative. This would be a data set almost four times the size of the current one, which would require a great deal of extra resource to acquire. By contrast, PCA allows us to observe significant effects using our relatively modest group size.

Conclusions

Taken together, our results represent a substantial body of evidence for microstructural changes during normal development, and their effect on general intelligence. As with all such studies, our findings are dependent on the specific age range of our participants, but we have seen that over this range of 8–16 yr, the development trajectories of white matter tracts differ significantly between the sexes, both in terms of the total volume of white matter and its diffusion characteristics. Whilst age has been

shown to be the biggest driver behind the latter, overall brain volume is also an important determining factor in FA and MD values. There is also a significant factor in both FA and MD that predicts IQ and which, given its secondary nature, would not be clear without the PCA-based approach taken here. Finally, we demonstrated evidence that a set of association tracts develop differently to the other white matter structures we studied. These findings will inform future imaging studies of normal development, and further understanding of the biophysical changes taking place in white matter as children grow into adults.

Funding

This work was supported by the UK Medical Research Council (programme grant number G0300117), and Engineering and Physical Sciences Research Council (grant number EP/C536851/1). CC wishes to acknowledge European Union Grant number FP7-ICT-2009-C 238292.

Acknowledgments

The authors are very grateful to all the participants and their families; and to Dr Martin King for useful discussions in relation to the statistical aspects of this work.

References

Asato M.R., Terwilliger R., Woo J. & Luna B. (2010). White matter development in adolescence: A DTI study. *Cerebral Cortex* **20**(9):2122–2131.

Ashtari M., Cervellione K.L., Hasan K.M., Wu J., McIlree C., Kester H., Ardekani B.A., Roofeh D., Szeszko P.R. & Kumra S. (2007). White matter development during late adolescence in healthy males: a cross-sectional diffusion tensor imaging study. *NeuroImage* **35**(2):501–510.

Barnea-Goraly N., Menon V., Eckert M., Tamm L., Bammer R., Karchemskiy A., Dant C.C. & Reiss A.L. (2005). White matter development during childhood and adolescence: a cross-sectional diffusion tensor imaging study. *Cerebral Cortex* **15**(12):1848–1854.

Basser P.J. & Pierpaoli C. (1996). Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance, Series B* **111**(3):209–219.

Behrens T.E.J., Johansen-Berg H., Jbabdi S., Rushworth M.F.S. & Woolrich M.W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage* **34**(1):144–155.

Benes F.M., Turtle M., Khan Y. & Farol P. (1994). Myelination of a key relay zone in the hippocampal formation occurs in the human brain during childhood, adolescence, and adulthood. *Archives of General Psychiatry* **51**(6):477–484.

Blanton R.E., Levitt J.G., Peterson J.R., Fadale D., Sporty M.L., Lee M., To D., Mormino E.C., Thompson P.M., McCracken J.T. & Toga A.W. (2004). Gender differences in the left inferior frontal gyrus in normal children. *NeuroImage* **22**(2):626–636.

Bonekamp D., Nagae L.M., Degaonkar M., Matson M., Abdalla W.M.A., Barker P.B., Mori S. & Horská A. (2007). Diffusion tensor imaging in children and adolescents: reproducibility, hemispheric, and age-related differences. *NeuroImage* **34**(2):733–742.

Bratchell N. (1989). Multivariate response surface modelling by principal components analysis. *Journal of Chemometrics* **3**(4):579–588.

Caprihan A., Pearlson G.D. & Calhoun V.D. (2008). Application of principal component analysis to distinguish patients with schizophrenia from healthy controls based on fractional anisotropy measurements. *NeuroImage* **42**(2):675–682.

Caviness V.S., Kennedy D.N., Richelme C., Rademacher J. & Filipek P.A. (1996). The human brain age 7-11 years: a volumetric analysis based on magnetic resonance images. *Cerebral Cortex* **6**(5):726–736.

Cerghet M., Skoff R.P., Bessert D., Zhang Z., Mullins C. & Ghandour M.S. (2006). Proliferation and death of oligodendrocytes and myelin proteins are differentially regulated in male and female rodents. *The Journal of Neuroscience* **26**(5):1439–1447.

Chiang M.C., Barysheva M., Shattuck D.W., Lee A.D., Madsen S.K., Avedissian C., Klunder A.D., Toga A.W., McMahon K.L., de Zubicaray G.I., Wright M.J., Srivastava A., Balov N. & Thompson P.M. (2009). Genetics of brain fiber architecture and intellectual performance. *The Journal of Neuroscience* **29**(7):2212–2224.

Clayden J.D., King M.D. & Clark C.A. (2009a). Shape modelling for tract selection. In G.Z. Yang, D. Hawkes, D. Rueckert, A. Noble & C. Taylor (Eds.), *Medical Image Computing and Computer-Assisted Intervention*, vol. 5762 of *Lecture Notes in Computer Science*, pp. 150–157. Springer-Verlag.

Clayden J.D., Muñoz Maniega S., Storkey A.J., King M.D., Bastin M.E. & Clark C.A. (2011). TractoR: Magnetic resonance imaging and tractography with R. *Journal of Statistical Software* **in press**.

Clayden J.D., Storkey A.J., Muñoz Maniega S. & Bastin M.E. (2009b). Reproducibility of tract segmentation between sessions using an unsupervised modelling-based approach. *NeuroImage* **45**(2):377–385.

De Bellis M.D., Keshavan M.S., Beers S.R., Hall J., Frustaci K., Masalehdan A., Noll J. & Boring A.M. (2001). Sex differences in brain maturation during childhood and adolescence. *Cerebral Cortex* **11**(6):552–557.

Deary I.J., Bastin M.E., Pattie A., Clayden J.D., Whalley L.J., Starr J.M. & Wardlaw J.M. (2006). White matter integrity and cognition in childhood and old age. *Neurology* **66**(4):505–512.

Deary I.J., Penke L. & Johnson W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience* **11**(3):201–211.

Eluvathingal T.J., Hasan K.M., Kramer L., Fletcher J.M. & Ewing-Cobbs L. (2007). Quantitative diffusion tensor tractography of association and projection fibers in normally developing children and adolescents. *Cerebral Cortex* **17**(12):2760–2768.

Fagiolo G., Waldman A. & Hajnal J.V. (2008). A simple procedure to improve FMRIB Software Library Brain Extraction Tool performance. *British Journal of Radiology* **81**(963):250–251.

Giorgio A., Watkins K., Chadwick M., James S., Winmill L., Douaud G., Stefano N.D., Matthews P., Smith S., Johansen-Berg H. & James A. (2010). Longitudinal changes in grey and white matter during adolescence. *NeuroImage* **49**(1):94–103.

Groeschel S., Vollmer B., King M.D. & Connelly A. (2010). Developmental changes in cerebral grey and white matter volume from infancy to adulthood. *International Journal of Developmental Neuroscience* **28**(6):481–489.

Gur R.C., Turetsky B.I., Matsui M., Yan M., Bilker W., Hughett P. & Gur R.E. (1999). Sex differences in brain gray and white matter in healthy young adults: correlations with cognitive performance. *The Journal of Neuroscience* **19**(10):4065–4072.

Kim J.H. & Juraska J.M. (1997). Sex differences in the development of axon number in the splenium of the rat corpus callosum from postnatal day 15 through 60. *Developmental Brain Research* **102**(1):77–85.

Lebel C., Walker L., Leemans A., Phillips L. & Beaulieu C. (2008). Microstructural maturation of the human brain from childhood to adulthood. *NeuroImage* **40**(3):1044–1055.

Lenroot R.K. & Giedd J.N. (2010). Sex differences in the adolescent brain. *Brain and Cognition* **72**(1):46–55.

Leonard C.M., Towler S., Welcome S., Halderman L.K., Otto R., Eckert M.A. & Chiarello C. (2008). Size matters: cerebral volume influences sex differences in neuroanatomy. *Cerebral Cortex* **18**(12):2920–2931.

Lin F., Yu C., Jiang T., Li K., Zhu C., Zhu W., Qin W., Duan Y., Xuan Y., Sun H. & Chan P. (2006). Discriminative analysis of relapsing neuromyelitis optica and relapsing-remitting multiple sclerosis based on two-dimensional histogram from diffusion tensor imaging. *NeuroImage* **31**(2):543–549.

Mabbott D.J., Noseworthy M., Bouffet E., Laughlin S. & Rockel C. (2006). White matter growth as a mechanism of cognitive development in children. *NeuroImage* **33**(3):936–946.

Muñoz Maniega S., Bastin M.E., McIntosh A.M., Lawrie S.M. & Clayden J.D. (2008). Atlas-based reference tracts improve automatic white matter segmentation with neighbourhood tractography. In *Proceedings of the ISMRM 16th Scientific Meeting & Exhibition*, p. 3318. International Society for Magnetic Resonance in Medicine.

Muetzel R.L., Collins P.F., Mueller B.A., Schissel A.M., Lim K.O. & Luciana M. (2008). The development of corpus callosum microstructure and associations with bimanual task performance in healthy adolescents. *NeuroImage* **39**(4):1918–1925.

Nagy Z., Westerberg H. & Klingberg T. (2004). Maturation of white matter is associated with the development of cognitive functions during childhood. *Journal of Cognitive Neuroscience* **16**(7):1227–1233.

Paus T., Zijdenbos A., Worsley K., Collins D.L., Blumenthal J., Giedd J.N., Rapoport J.L. & Evans A. (1999). Structural maturation of neural pathways in children and adolescents: in vivo study. *Science* **283**(5409):1908–1911.

Penke L., Muñoz Maniega S., Murray C., Gow A.J., Valdés Hernández 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. *The Journal of Neuroscience* **30**(22):7569–7574.

R Development Core Team (2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Robinson E.C., Hammers A., Ericsson A., Edwards A.D. & Rueckert D. (2010). Identifying population differences in whole-brain structural networks: A machine learning approach. *NeuroImage* **50**(3):910–919.

Schmithorst V.J., Holland S.K. & Dardzinski B.J. (2008). Developmental differences in white matter architecture between boys and girls. *Human Brain Mapping* **29**(6):696–710.

Schmithorst V.J., Wilke M., Dardzinski B.J. & Holland S.K. (2002). Correlation of white matter diffusivity and anisotropy with age during childhood and adolescence: a cross-sectional diffusion-tensor MR imaging study. *Radiology* **222**(1):212–218.

Schmithorst V.J., Wilke M., Dardzinski B.J. & Holland S.K. (2005). Cognitive functions correlate with white matter architecture in a normal pediatric population: a diffusion tensor MRI study. *Human Brain Mapping* **26**(2):139–147.

Smith S.M. (2002). Fast robust automated brain extraction. *Human Brain Mapping* **17**(3):143–155.

Stacklies W., Redestig H., Scholz M., Walther D. & Selbig J. (2007). pcaMethods – A Bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23**(9):1164–1167.

Tamnes C.K., Østby Y., Walhovd K.B., Westlye L.T., Due-Tønnessen P. & Fjell A.M. (2010). Intellectual abilities and white matter microstructure in development: a diffusion tensor imaging study. *Human Brain Mapping* **31**(10):1609–1625.

Tau G.Z. & Peterson B.S. (2010). Normal development of brain circuits. *Neuropsychopharmacology* **35**(1):147–168.

Teipel S.J., Stahl R., Dietrich O., Schoenberg S.O., Perneczky R., Bokde A.L.W., Reiser M.F., Möller H.J. & Hampel H. (2007). Multivariate network analysis of fiber tract integrity in Alzheimer's disease. *NeuroImage* **34**(3):985–995.

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**(2):531–541.

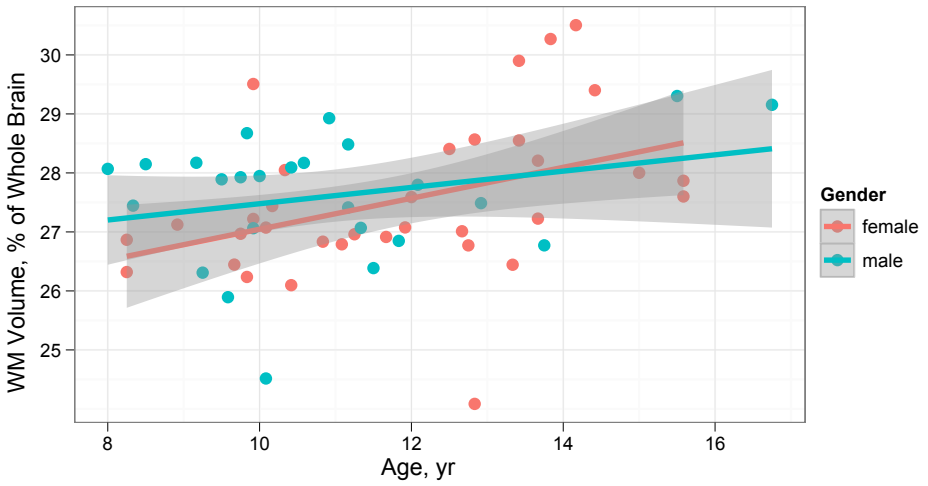
Westlye L.T., Walhovd K.B., Dale A.M., Bjornerud A., Due-Tønnessen P., Engvig A., Grydeland H., Tamnes C.K., Østby Y. & Fjell A.M. (2010). Life-span changes of the human brain white matter: Diffusion tensor imaging (DTI) and volumetry. *Cerebral Cortex* **20**(9):2055–2068.

Wickham H. (2009). *ggplot2: Elegant graphics for data analysis*. Springer, New York.

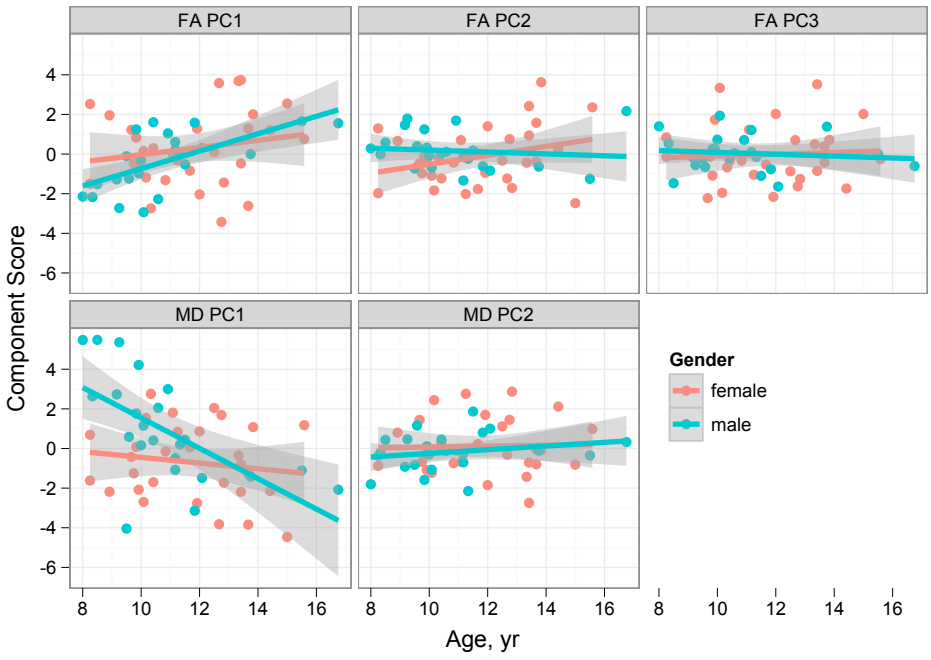
Wold H. (1966). Estimation of principal components and related models by iterative least squares. In P.R. Krishnaiaah (Ed.), *Multivariate Analysis*, pp. 391–420. Academic Press, New York.

Yakovlev P.I. & Lecours A. (1967). The myelogenetic cycles of regional maturation of the brain. In A. Minkowski (Ed.), *Regional development of the brain in early life*, pp. 3–70. Blackwell, Oxford.

Zhang Y., Brady M. & Smith S.M. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging* **20**(1):45–57.



Supplementary Figure 1: Scatter plot of age against white matter volume, expressed as a percentage of total brain volume.



Supplementary Figure 2: Scatter plots of age against component score for each of the principal components with greater than 10% of the variance. Linear regression lines and associated standard errors are shown for each gender.