Department of Computer Science, University of Otago



Te Whare Wānanga o Otāgo

Technical Report OUCS-2013-12

Functional network analysis of aging and Alzheimer's Disease: Results

Authors:

Paul McCarthy, Lubica Benuskova Department of Computer Science, University of Otago, New Zealand

Elizabeth Franz Department of Psychology, University of Otago, New Zealand

Note: The associated data files may be used for research purposes. Please cite this report.



Department of Computer Science, University of Otago, PO Box 56, Dunedin, Otago, New Zealand

http://www.cs.otago.ac.nz/research/techreports.php

Functional network analysis of aging and Alzheimer's Disease: Results

Paul McCarthy^{*} Lubica Benuskova[†] Elizabeth Franz[‡] University of Otago, Dunedin, New Zealand

October 30, 2013

Abstract

This report, and the associated data files, present the results of a functional network analysis conducted upon a fMRI data set, with the aim of identifying differences, in various functional network properties, between a group of healthy young individuals, a group of healthy aged individuals, and a group of individuals diagnosed with probable to mild Alzheimer's Disease. This report is intended as a complement to the PhD thesis Functional network analysis of aging and Alzheimer's Disease, by Paul McCarthy, submitted for the degree of Doctor of Philosophy, at the University of Otago, Dunedin, New Zealand [12].

Contents

1	Met	thods
	1.1	Preprocessing
	1.2	Data preparation
	1.3	Functional network creation
	1.4	Network measure calculation
	1.5	Statistical analysis
2	Res	ults
	2.1	Functional networks
	2.2	Functional network data
	2.3	Connectivity analysis
	2.4	Global analysis
	2.5	Regional analysis
	2.6	Voxelwise analysis

Overview

Section 1 describes in detail the methods used for data preprocessing and preparation, functional network creation and statistical analysis. Section 2 describes the data files that accompany this report, which comprise the following:

File name	Size	$MD5 \ checksum$	Description
networks_young.tbz2	2.5 GB	c9eaae4d26fe5ed8c130d1f2f4ae6149	Functional networks for every subject in the young group.
networks_awout.tbz2	2.5 GB	4ef2a7125f6a9c512ead655cfea995a8	Functional networks for every subject in the aged group.
networks_awith.tbz2	2.2 GB	e5c2e1962d6cc50b21a6f140258b6419	Functional networks for every subject in the aged with AD group.
network_data_young.tbz2	576 MB	8bdc5526e319cdcc1b800a613999b61e	Functional network data for every subject in the young group.
network_data_awout.tbz2	551 MB	e7081971f19a70e77ca3550d92f2c97d	Functional network data for every subject in the aged group.
network_data_awith.tbz2	473 MB	914eba180404aaa250a5044c24dfcfa9	Functional network data for every subject in the aged with AD group
connectivity_analysis.tbz2	3.8 MB	b6dcafdeafddc1e750b4e555e8169c1f	Results for the statistical analysis of regional connectivity.
$global_analysis.tbz2$	535 KB	4953b57874be33892b0b8b19d6f8338a	Results for the statistical analysis of global network measures.
regional_analysis.tbz2	7.3 MB	4c805e84fe26dd57b9fcf289b3a5a06a	Results for the statistical analysis of regional network measures.
$voxelwise_analysis.tbz2$	492 MB	6e293939e44b41714140bead07990679	Results for the statistical analysis of voxelwise network measures.

^{*}pauldmccarthy@gmail.com; corresponding author

[†]lubica@cs.otago.ac.nz

[‡]lfranz@psy.otago.ac.nz

1 Methods

We performed voxelwise functional network analysis upon fMRI data originally collected by Buckner *et al.* [3] [data set #2-2000-118W, 5]. Structural and functional MRI data were acquired from 41 subjects: 14 young (9 females/5 males, mean age 21.1, SD 2.0), 15 healthy aged subjects (9 females/6 males, mean age 75.1, SD 6.9), and 12 aged subjects (7 females/5 males, mean age 77.1, SD 5.3) who had been clinically diagnosed with Dementia of the Alzheimer Type. These three groups shall be referred to as **young**, **aged**, and **aged with AD** respectively. The individuals in the aged and aged with AD groups were clinically assessed for the presence of dementia using the Clinical Dementia Rating [CDR, 9], with all individuals in the aged group scoring CDR 0. Of the 12 individuals in the aged with AD group, seven scored CDR 0.5, corresponding to a diagnosis of probable AD; the remaining five scored CDR 1, corresponding to a diagnosis of mild AD.

Structural MRI images consisted of 128 1.25mm slices, with each slice containing $256 \times 256 \ 1mm^2$ in-plane isotropic voxels; these images were resampled to $128 \times 128 \times 75 \ 2mm^3$ isotropic voxels. Each functional MRI image consisted of 16 8mm slices, acquired parallel to the anterior-posterior commissure, with each slice consisting of $64 \times 64 \ 3.75mm^2$ in-plane isotropic voxels. Functional image slices were acquired in an interleaved manner, from superior to inferior, with even slices acquired first¹ [17]. Functional image acquisition (TR) time was 2.68 seconds.

The study involved subjects completing a simple visual motor task within an event based experimental paradigm. Each subject underwent four fMRI recording sessions²; within each session, 128 fMRI images were acquired over a period of 5 minutes, 43 seconds. During a single session, 15 trials were executed, with each trial consisting of either one or two visual stimuli, a checkerboard pattern flickering at 8Hz, displayed for 1.5 seconds. The subjects were instructed to push a button with their right index finger upon onset of each stimulus. During a 'one-stimulus' trial, the stimulus was triggered at the start of the trial. During a 'two-stimulus' trial, the first stimulus was triggered at the start of the trial bad a duration of 21.44 seconds (8 image acquisitions), the first trial in each session began 10.72 seconds after the beginning of image acquisition (at image #5), and the last trial ended at 5 minutes, 32 seconds (at image #124).

Some discrepancies regarding the classification of healthy aged subjects, and subjects diagnosed with AD, are apparent in the original study. One subject was listed as having scored CDR 0 in the assessment for DAT, but was subsequently placed in the aged with AD group. Another subject scored CDR 0, and was correctly placed in the aged group, but performed very poorly during the experiment in both reaction time and misses³. For this analysis, we moved the former subject from the aged with AD group to the aged group, but left the latter subject in the aged group.

The original study made use of a simple sensorimotor task, which allowed us to explore age and AD related changes in both sensorimotor activity, and in resting state activity. We accomplished this by splitting our analysis into two parts. In the first part, we analysed fMRI data created by taking the average of every trial for each subject, in a similar manner to the analysis of Buckner *et al.* [3]. This analysis focuses upon the task related connectivity which was present within a single one- or two-stimulus trial, and is hence referred to as the **task based** analysis⁴. The second part of the analysis, instead of focusing on task related connectivity, focuses upon functional connectivity which persisted throughout the duration of the experiment, thus reflecting background, or resting state connectivity. This part is hence referred to as the **resting state** analysis, and was made possible by concatenating the data acquired for each session into a single long fMRI volume. Justification for the validity of this approach is provided by Greicius *et al.* [6] who used the same technique upon the very same data set to identify AD related changes in default mode network activity.

There were five stages to the analysis, each of which shall be described in detail in this section. First, the data were preprocessed to remove noise and enable comparison across subjects and groups. Next, the data were prepared for both the task based and resting state analyses. Then a range of functional networks were created, and network measures calculated upon them. Finally, a statistical analysis was carried out, comparing the three groups with each other. For brevity, the details in the following sections (with the exception of the statistical analysis details in Section 1.5) pertain to a single subject; however, the same processing and analysis steps were performed on the data for all subjects.

1.1 Preprocessing

Preprocessing applied to the four fMRI volumes (one volume per session) consisted of the steps described below. In each session, the first trial began at acquisition of fMRI image #5, and the last trial ended at acquisition of image #124. Therefore, before any preprocessing, the first four and last four images from every session were discarded.

¹If each slice is given an index from 1 to 16, with 1 being the most inferior slice and 16 the most superior, the slice acquisition order was [16, 14, 12, 10, 8, 6, 4, 2, 15, 13, 11, 9, 7, 5, 3, 1].

 $^{^{2}}$ One subject (from the aged with AD group) only underwent three sessions, for unknown reasons.

 $^{{}^{3}}A$ miss refers to the case where the subject did not push the button in response to the visual stimulus.

 $^{^{4}}$ We decided that there was no need to discriminate between one- and two-stimulus trials, after conducting a GLM analysis upon the task based data and comparing the results to the original results of Buckner *et al.* [3]; [see 12, Appendix A].

- 1. Visual inspection: Every volume was visually inspected to check for obvious anomalies. This step uncovered three suspect data sets, all from the aged group. The data for two subjects contained aliasing effects, which were manually corrected. The data for one other subject were discarded, due to the presence of significant noise throughout every session. Thus, the number of individuals in the aged group was reduced to 14 (9 females/5 males, mean age 74.9, SD 6.9).
- 2. *Image padding* All fMRI and MRI images were padded with empty voxels (i.e. voxels with a value of 0); this was done to prevent image cropping during any of the image transformation steps.
- 3. *Slice-timing correction*: Every volume was corrected for slice timing differences using Fourier interpolation; this was accomplished with the 3dTshift tool, provided with AFNI [4].
- 4. High-pass temporal filtering: To correct for scanner drift, a high-pass temporal filter was applied to every volume, with a pass frequency of $1/42.88 \approx 0.02$ Hz (the duration of two consecutive trials). This was achieved using the fslmaths utility, provided with FSL [16].
- 5. Motion correction: Motion correction was applied to each volume using the mcflirt utility, provided with FSL [10].
- 6. Intra-session alignment To allow for averaging across sessions, the fMRI images from every session were aligned to the first image of the first session, using FSL's flirt utility to calculate and apply a 6 parameter rigid body transformation.
- 7. Intensity normalisation: To allow for concatenation of session data for the resting state analysis, the baseline intensity of every volume was shifted to zero, by subtracting the temporal mean at each voxel.

Further preprocessing steps to the structural MRI data, and calculation of transformations between fMRI images, MRI images, and the MNI152 T1 standard template [11], were as follows:

- 8. Brain segmentation: Non-brain matter (e.g. skull tissue) was removed from the structural MRI image using the bse utility provided with BrainSuite [14, 15].
- 9. *Tissue classification*: The structural MRI image was corrected for bias field inhomogeneities, and each MRI voxel classified as white matter, grey matter, or cerebrospinal fluid, using BrainSuite's bfc and pvc utilites.
- 10. Intra-subject registration: A mean fMRI image was created from every fMRI volume, and a 12 parameter affine transformation to the subject's structural MRI image calculated, as implemented in FSL's flirt utility.
- 11. Spatial normalisation: A 12 parameter affine transformation from the structural MRI image to the 2mm³ MNI152 T1 standard brain template was calculated using FSL's flirt; this transformation was then used to initialise a non-linear transformation using the fnirt utility provided with FSL [1].

Spatial smoothing of the fMRI time series images was explicitly not applied, to prevent the introduction of artificial correlations between adjacent voxels [7, 19]. Furthermore, we decided that spatial smoothing for the purposes of improving cross-subject anatomical alignment was unnecessary, due to the smoothing which is inherent in the transformation of fMRI and network measure images to MNI152 space.

1.2 Data preparation

From the four session fMRI volumes, two fMRI volumes were created for analysis: one task based volume of 21.44 seconds duration (8 image acquisitions), and one resting state volume 21 minutes, 26.4 seconds (480 image acquisitions) duration.

A binary mask image was created to ensure that non-brain voxels were excluded from analysis. This restriction was accomplished by using the inverse of the MRI to fMRI transformation, calculated in preprocessing step 10, to transform the tissue classification image, generated in step 9, into the subject's native space; this image was then used to mask all fMRI voxels that had not been classified as white matter, grey matter, or cerebrospinal fluid.

A similar process was used to anatomically label each voxel in fMRI space. This was performed to enable analysis of regional connectivity and regional network measures, and was accomplished by transforming the anatomical atlas from standard space into fMRI space. The anatomical atlas used was a reduced version of the LPBA40 atlas [11], and is summarised in Table 1.

In order to explore the effects of spatial normalisation upon functional networks, copies of both the task based and resting state volumes were transformed from the subject's native space into MNI152 space⁵. These MNI152 space volumes were then resampled and interpolated back to the original native space resolution of $64 \times 64 \times 16$ voxels; refer to Table 2 for details of the voxel dimensions of native and MNI152 space networks. Functional network analysis was then performed on all four analysis volumes: the task based volumes in both fMRI and standard space, and the resting state volumes in both spaces.

 $^{^{5}}$ The terms *native space* and *MNI152 space* shall be used interchangeably with *fMRI space* and *standard space*, respectively.

$An atomical\ label$	Code	$Corresponding \ LPBA40 \ labels$	Code
L/R Frontal	311/301	L/R superior frontal gyrus	21/22
		L/R middle frontal gyrus	23/24
		L/R inferior frontal gyrus	25/26
		L/R precentral gyrus	27/28
L/R Orbito-Frontal	312/302	L/R middle orbitofrontal gyrus	29/30
		L/R lateral orbitofrontal gyrus	31/32
		L/R gyrus rectus	33/34
L/R Parietal	313/303	L/R postcentral gyrus	41/42
		L/R superior parietal gyrus	43/44
		L/R supramarginal gyrus	45/46
		L/R angular gyrus	47/48
		L/R precuneus	49/50
L/R Occipital	314/304	L/R superior occipital gyrus	61/62
		L/R middle occipital gyrus	63/64
		L/R inferior occipital gyrus	65/66
		L/R cuneus	67/68
L/R Temporal	315/305	L/R superior temporal gyrus	81/82
		L/R middle temporal gyrus	83/84
		L/R inferior temporal gyrus	85/86
		L/R fusiform gyrus	91/92
		L/R insular cortex	101/102
L/R Limbic	316/306	L/R parahippocampal gyrus	87/88
		L/R lingual gyrus	89/90
		L/R cingulate gyrus	121/122
		L/R hippocampus	165/166
L/R Basal Ganglia	317/307	L/R caudate	161/162
		L/R putamen	163/164
L/R Cerebellum	318/308	L/R cerebellum	181
Brainstem	309	Brainstem	182
No label	300	No label	0

Table 1: Anatomical labels and codes used, and mapping to the LPBA40 atlas.

Table 2: Dimensions of native/fMRI and MNI152/standard space images and functional networks. Both fMRI and standard space images adhere to radiological, or LAS orientation, where: x denotes the location along the axis orthogonal to the sagittal plane, and increases from right to left; y denotes the location along the axis orthogonal to the coronal plane, and increases from posterior to anterior; and z denotes the location along the axis orthogonal to the transverse plane, and increases from inferior to superior. When voxel/node locations are specified in millimetres, the origin (i.e. the node/voxel at location x = 0, y = 0, z = 0) in native/fMRI space networks is the most right, posterior and inferior voxel. However, the origin in standard/MNI152 space millimetre coordinates is at the anterior-posterior commisure.

Space	Nur	Number of voxels		Voxel length (mm)		nm)
	x	y	z	x	y	z
fMRI Standard	$\begin{array}{c} 64 \\ 64 \end{array}$	$\begin{array}{c} 64 \\ 64 \end{array}$	16 16	$3.75 \\ 2.84375$	$3.75 \\ 3.40625$	$8.0 \\ 11.375$

1.3 Functional network creation

A correlation matrix was created for each analysis volume, by calculating the Pearson correlation coefficient between all pairs of included voxels x and y:

$$r(x,y) = \frac{\sum (x_i - \overline{x}) (y_i - \overline{y})}{\sqrt{\sum (x_i - \overline{x})^2 \sum (y_i - \overline{y})^2}}$$
(1)

where x_i and y_i is the BOLD contrast in voxels x and y at time i, respectively, and \overline{x} and \overline{y} are the means across time for voxels x and y respectively. Three correlation thresholds were used to create unweighted and undirected functional networks for each of the four correlation matrices, as listed in Table 3. The thresholds were empirically selected to ensure that the resulting networks were of a density suitable for analysis.

Three additional sub-networks were derived from these whole brain functional networks. The MRI tissue classification images, in both standard and fMRI space, were used to extract networks consisting of only white matter voxels only grey matter voxels, and only white and grey matter voxels (i.e. with cerebrospinal fluid voxels removed). Finally, two additional networks were extracted from each of these networks, one consisting of only positive correlations, and one consisting of negative correlations.

To summarise, correlation matrices were generated for the resting state and task based fMRI volumes in both fMRI and standard space. Three unweighted and undirected functional networks were created from each correlation matrix over a range of correlation thresholds. Four sub-networks were derived from each thresholded network: whole brain, white+grey matter, grey matter and white matter. These networks were further subdivided into three networks, one with edges representing absolute correlations which passed the threshold, one with edges representing positive correlations only, and one with edges representing negative correlations. This process resulted in a total of 144 classes of functional networks, as listed in Table 4.

Table 3: Correlation thresholds used for creation of functional networks from both the task based and resting state correlation matrices. The same thresholds were used for both fMRI space and standard space matrices. The *n* column lists the number of time points in the calculation of the correlation coefficient, the r_c column lists the correlation thresholds that were used, and the α column lists the corresponding two-tailed edge level significance (uncorrected).

	n	r_c	α
Task based	8	$0.818398 \\ 0.899876 \\ 0.962249$	1×10^{-2} 1×10^{-3} 1×10^{-5}
Resting state	480	$0.446347 \\ 0.534151 \\ 0.629034$	1×10^{-25} 1×10^{-38} 1×10^{-58}

Table 4: For each subject, functional networks were created for all combinations of data set, space, correlation threshold, tissue type, and threshold type. This resulted in 144 functional networks for each subject, and a total of 5760 networks across all 40 subjects.

Data set	Space	Threshold	Tissue type	Threshold type
Task based Resting state	fMRI space Standard space	Refer to Table 3	Whole brain White+grey matter Grey matter White matter	Absolute Positive Negative

The purpose of analysing white matter+grey matter, white matter and grey matter networks, alongside the whole brain networks from which they were derived, is to identify the source of any group differences which are found in the statistical analysis. For example, a group difference which is present in whole brain, white+grey matter and grey matter networks, but not in white matter networks, is likely to be localised to grey matter networks. Conversely, a group difference which is present in whole brain networks but not in white matter+grey matter networks is likely to be caused by interactions within BOLD signals originating within voxels representing cerebro-spinal fluid. A similar line of reasoning is behind the approach of performing analysis upon positively and negatively correlated networks, an alongside the absolute correlation networks from which they were derived. This allows the attribution of any observed group differences to either positive or negative connectivity.

One final step was performed prior to the calculation of network measures: disconnected nodes and small isolated components were removed from each network. This process is referred to as *pruning*, and was performed to prevent biasing of global, regional and voxelwise network measures. For each network, any disconnected nodes were removed. Then, if the remaining network consisted of one major component which constituted at least 75% of the total network size, all other minor components were removed. For more fragmented networks consisting of multiple components, with no single major component, only the disconnected nodes were removed. This had the effect that, for these fragmented networks, some network measures could not be calculated. These fragmented networks were therefore excluded from statistical analysis of those measures.

1.4 Network measure calculation

Every class of functional network (see Table 4) was analysed at three 'resolution' levels: *global*, *regional*, and *voxelwise*. Functional connectivity within and between regions was also analysed. Global analysis involved the calculation of measures at the network level, such as density and mean path length. Regional analysis involved the calculation of network measures by anatomical region. This typically involved calculating the mean voxelwise measure across every node within each region, but also included regionally specific measures such as size (number of nodes) and regional efficiency. Voxelwise analysis involved the calculation and comparison of node measures at each voxel in the brain, such as degree and local efficiency. Finally, functional connectivity was calculated as the intra- and inter-regional densities within and between all regions in each network. All network measures included in the analyses are listed in Table 5.

Some important points must be noted. The global and regional numbers of pruned nodes represent the disconnected nodes and small components which were removed before further analysis. The global and regional numbers of nodes reflects the network and region sizes *before* these disconnected nodes were removed, and hence solely represent global and regional brain volumes. Mean network degree was not calculated at the global level, as it is essentially a proxy for network density. Instead, the maximum degree of the network, the degree of the most highly connected node, was calculated. Mean degree and node degree were, however, calculated as part of the regional and voxelwise analyses respectively. Global efficiency and mean local efficiency were calculated at the network level, regional efficiency at the regional level, and local efficiency at the voxel level. Fragmented networks (i.e. networks containing multiple components) were excluded from global, regional and voxelwise analysis of path length, clustering coefficient, and the small world index.

Table 5: The network measures included in each stage of the analysis.

Measure	Global	Regional	Voxelwise	Connectivity
Number of nodes Number of pruned nodes Density Degree Clustering coefficient Path length Local efficiency Edge distance Small world index Global efficiency Assortativity Number of components	$\begin{array}{c} \checkmark^{1} \\ \checkmark \\ \checkmark \\ \checkmark^{4} \\ \checkmark \\ $	$\begin{array}{c} \checkmark^{1} \\ \checkmark \\ \checkmark^{2} \\ \checkmark \\ \checkmark \\ \checkmark \\ \checkmark \\ \checkmark^{5} \\ \checkmark^{6} \end{array}$	$ \begin{array}{c} \checkmark \\ \checkmark $	\checkmark^3

¹ Before pruning of disconnected nodes/small components.

 2 As part of the connectivity analysis.

³ Intra- and inter-regional density.

⁴ Maximum degree.

⁵ Regional efficiency.

⁶ Mean and voxelwise local small world index.

⁷ Only for those network classes that contained fragmented networks.

1.5 Statistical analysis

Properties of all classes of functional networks were compared across each of the young, aged, and aged with AD groups, in order to find differences related to age and the presence of AD. Statistical analysis was performed separately for each of the three resolution levels (global, regional, and voxelwise), and for regional connectivity, against the null hypothesis of no difference, in any functional network properties, between each pair of groups. Where a fragmented network was exluded from analysis of a specific network measure, the degrees of freedom used in statistical calculation was adjusted accordingly. For each subject, regional and voxelwise network measures were normalised to Z scores in order to account for the effects of varying network size and density across subjects [2]. However, to assess the effects of this normalisation step, both normalised and unnormalised data underwent statistical analysis. For every functional network class, and every pair of groups, the statistical analysis proceeded as follows.

Global analysis For each global network measure, the Mann-Whitney U test was used to test for any difference between each pair of groups. A two-tailed likelihood of $\alpha = 0.05$ was used to test for significance.

Regional analysis Network measures, standardised to Z scores for each subject, were compared within each region, as defined by the reduced LPBA40 atlas. This meant that multiple tests, one for each region, were performed for each network measure between the groups. Therefore, multiple comparisons correction was incorporated into the testing using a nonparametric permutation approach [13]. For each pair of groups, Student's t test was used to calculate a t value at each region. Then, the group labels were randomly permuted 10000 times, and t values calculated at each region for every permutation. In this way a distribution of the maximal t value, across all regions, was generated; use of the maximal t value ensured control of the family wise Type I error rate. The real t values were then compared against this maximal t distribution to determine their likelihood of occurrence, again using a two-tailed likelihood of $\alpha = 0.05$.

Connectivity analysis Intra- and inter-regional densities were compared between each pair of groups, in much the same way as for the regional analysis. Student's t test was used to calculate a t value for every region (intra-regional density), and for every pair of regions (inter-regional density). A maximal t distribution was generated from 10000 random permutations of group labels, and the real t values compared against this distribution using a two-tailed likelhood of $\alpha = 0.05$ to determine significance.

Voxelwise analysis Voxelwise network measure images, standardised to Z scores for each subject, were compared between each pair of groups using nonparametric cluster size thresholding [8]. fMRI space images were transformed to MNI152 space, and masked using each subject's binary mask image (also transformed to MNI152 space) before voxelewise analysis proceeded. This transformation step was not necessary for standard space images. Student's t test was used to create an image of t values upon the network measure images from each pair of groups. This t image was subjected to cluster size thresholding using a t threshold of 3.0. Then, t images were generated from 10000 random permutations of group labels, and a maximal cluster size distribution created. The thresholded clusters in the observed t image were declared as significant if their size was in the top 95th percentile of the maximal cluster size distribution (i.e. a cluster level significance of $\alpha = 0.05$). For a voxel to be included in a statistical test between two pairs of groups, a value had to be present at that voxel for at least 90% of all subjects from both groups. In other words, if more than 10% of subjects were missing a value at a particular voxel, that voxel was excluded from analysis.

In order to overcome the problem of missing voxels, either due to their representing disconnected nodes, or due to subject specific atrophy, values at missing voxels were imputed using mean replacement by random neighbour selection, an adaptation of techniques described by Vaden Jr *et al.* [18]. This imputation process was applied before the statistical test decsribed above. For a missing value in a subject's network measure image, a replacement value was generated by randomly selecting the values of up to 5 voxels from the images of all other subjects in the same group, located within a sphere of radius 10mm, centred at the missing voxel. The mean of these sampled voxels was used as the replacement value. Imputation only proceeded if 90% of the randomly sampled voxels were present (i.e. the sampled voxels were not also missing values). After imputation, any missing voxels which could not be imputed were simply left as 'missing'. In order to assess the effect of imputation, both imputed and non-imputed voxelwise data underwent statistical analysis.

2 Results

The data files that accompany this report are listed in Table 6, and include data and results for the global, regional and node level network data, for every functional network which was created as part of the analysis. Each file has been compressed using the UNIX tar utility, with the command tar cjf [filename].tbz2 [directory], and will therefore need to be decompressed with a command such as tar xjf [filename.tbz2]. Due to size constraints, structural MRI images and functional MRI time series data have not been provided. The raw functional correlation matrices have not been provided for the same reason. Please contact the corresponding author if you require access to the MRI and fMRI data, and/or to correlation matrices.

Table 6: Data files which accompany this report.

File name	Description
networks_young.tbz2	Functional networks for every subject in the young group.
networks_awout.tbz2	Functional networks for every subject in the aged group.
networks_awith.tbz2	Functional networks for every subject in the aged with AD group.
network_data_young.tbz2	Functional network data for every subject in the young group.
network_data_awout.tbz2	Functional network data for every subject in the aged group.
network_data_awith.tbz2	Functional network data for every subject in the aged with AD group.
connectivity_analysis.tbz2	Results for the statistical analysis of regional connectivity.
global_analysis.tbz2	Results for the statistical analysis of global network measures.
regional_analysis.tbz2	Results for the statistical analysis of regional network measures.
voxelwise_analysis.tbz2	Results for the statistical analysis of voxelwise network measures.

Throughout this section, and throughout the associated data files, the conventions listed in Table 7 are adhered to, when referring to data or results from specific groups, data sets, network classes and measures.

Table 7: Naming conventions used for data sets, network and threshold types, and network measures.

Group	Young Aged	young awout
	Aged with AD	awith
Dataset	Task based Resting state	trialavg concat
Space	fMRI/native space MNI152/standard space	fmri std
Network type	Whole brain White+grey matter Grey matter White matter	whole nocsf gm wm
Threshold type	Absolute Positive Negative	abs pos neg
Measure	Number of nodes after pruning Number of nodes before pruning Number of pruned nodes Regional size Number of components Size of the largest component Density Degree Maximum degree Assortativity Closeness centrality Closeness centrality Clustering coefficient Edge distance Global efficiency Local efficiency Path length Small world index Local small world index	nodes orignodes pruned size edges components largest_component density degree max_degree assortativity closeness clustering edgedist global_efficiency local_efficiency efficiency pathlength smallworld_index

2.1 Functional networks

The files which have a name of the form networks_*.tbz2 contain the functional networks for every subject and network class⁶. Once all of the files have been decompressed, the data is organised as follows:



where group is one of young, awout or awith, and subject_no is a two-digit identifier for every subject in the analysis (e.g. 07/). Within the fmri and std sub-directories are plain text files which describe every functional network that was included in the analysis. Each network is described in two files, which are named according to the following convention:

[network-type]_[threshold-direction]_[threshold]_nodes.txt [network-type]_[threshold-direction]_[threshold]_edges.txt

where:

- [network-type] is one of whole, nocsf, gm, or wm.
- [threshold-direction] is one of abs, pos, or neg.
- [threshold] is one of 0.818398, 0.899876, or 0.962249 for task based networks, or 0.446347, 0.534151, or 0.629034 for resting state networks.

Files which end with nodes.txt define the nodes in the network, as shown in Figure 1. Each line describes one node in the network, and contains the node ID, x, y, and z coordinates, and anatomical labels. The first anatomical label (e.g. 318) refers to the region in the reduced LPBA40 atlas; the second label (e.g. 181) refers to the region in the full LPBA40 atlas; and the third label refers to the tissue class (CSF = 1, GM = 2, WM = 3), as generated by the Brainsuite pvc utility, in Preprocessing step 9 (Section 1.1). Refer to Table 2 for details regarding node/voxel dimensions, and to Table 1 for details regarding anatomical labels.

0	135.000000,41.250000,0.000000 318 181 2
1	138.750000,41.250000,0.000000 318 181 2
2	90.000000,45.000000,0.000000 308 180 2
3	93.750000,45.000000,0.000000 308 180 1
·	

Figure 1: Example network nodes.txt file.

Files ending in edges.txt define the edges between nodes in the network; an example is given in Figure 2. Each line defines a single unweighted, undirected edge between two nodes, with the two values referring to node IDs, as listed in the corresponding nodes.txt file.

0	1	
0	21	
0	26	
0	65	
0	102	
.		
.		

Figure 2: Example network edges.txt file.

⁶Only the data for pruned functional networks, i.e. after disconnected nodes and small components were removed, have been provided. Please contact the corresponding author if you require access to the unpruned networks.

2.2 Functional network data

The files network_data_young.tbz2, network_data_awout.tbz2, and network_data_awith.tbz2 contain global, regional and node level network measure data for every functional network which was created as part of the analysis. The data (once all three files have been decompressed) is organised as follows:



where group is one of young, awout or awith, and subject_no is a two-digit identifier for every subject in the analysis (e.g. 07/). Files containing network data, for each subject, are located within the fmri/ and std/ sub-directories. Within these sub-directories, eight data files are present for each network type, threshold direction and threshold, and are named according to the following convention:

```
[network-type]_[threshold-direction]_[threshold]_global.txt
[network-type]_[threshold-direction]_[threshold]_node.txt
[network-type]_[threshold-direction]_[threshold]_regional.txt
[network-type]_[threshold-direction]_[threshold]_connectivity.txt
[network-type]_[threshold-direction]_[threshold]_prune_global.txt
[network-type]_[threshold-direction]_[threshold]_prune_node.txt
[network-type]_[threshold-direction]_[threshold]_prune_regional.txt
[network-type]_[threshold-direction]_[threshold]_prune_regional.txt
```

where:

- [network-type] is one of whole, nocsf, gm, or wm.
- [threshold-direction] is one of abs, pos, or neg.
- [threshold] is one of 0.818398, 0.899876, or 0.962249 for task based networks, or 0.446347, 0.534151, or 0.629034 for resting state networks.

Files which have **prune** in their name contain data from networks after disconnected nodes and isolated components were removed. Conversely, files which do not have **prune** in their name contain data from networks before disconnected nodes and isolated components were removed.

Files ending in global.txt contain global network measures, with one measure per line, as space separated name-value pairs. An example is given in Figure 3.

```
nodes 9356.000000
edges 218657.000000
density 0.004996
degree 46.742000
max_degree 707.000000
components 3462.000000
largest_component 5820.000000
connected 5961.000000
disconnected 3395.000000
clustering 0.380311
pathlength 2.232399
smallworld_index 93.021256
global_efficiency 0.121579
local_efficiency 0.391598
assortativity 0.180532
```

Figure 3: Example global.txt file.

Files ending in node.txt contain node (or voxel) level network measure data, with one line for every node in each network. As depicted in Figure 4, the first line of a node.txt file contains space separated column headers, with each subsequent line containing space separated values for the respective network measures. The x, y, and z columns give the voxel coordinates of each node in the network (refer to Table 2 for details regarding node/voxel dimensions). The node column gives the node ID for each node, and the label column gives anatomical labels for each node, using the codes listed in Table 1. The component column gives the index of the component in which each node is a part. Component

node x y z label degree clustering local_efficiency pathlength closeness edgedist local_smallworld_index component 0.000000 28.000000 14.000000 0.000000 308.000000 35.000000 0.847059 0.923529 3.137653 0.318710 17.308374 49.640423 0.000000 1.000000 40.000000 14.000000 0.000000 318.000000 1.000000 1.000000 3.470356 0.288155 11.710431 52.984977 0.000000 2.000000 33.000000 19.000000 0.000000 309.000000 256.000000 0.593903 0.795629 2.795498 0.357718 12.095098 39.064592 0.000000 3.000000 34.000000 19.000000 0.000000 309.000000 6.000000 1.000000 3.450937 0.289776 11.246026 53.283135 0.000000 .

Figure 4: Example node.txt file.

indices begin at 0 - if a network is connected, the component value for every node will be 0. Refer to Table 7 for the meanings of the other columns.

Files ending in regional.txt contain network measures calculated upon the nodes in each of the anatomical regions (refer to Table 1). An example regional.txt file is shown in Figure 5.

region size degree efficiency clustering closeness local_smallworld_index pathlength edgedist 300 255.000000 391.760800 0.290358 0.481483 0.330571 4.886922 3.115430 16.432773 301 777.000000 264.249700 0.275748 0.436921 0.310498 4.190995 3.315125 21.384746 302 161.000000 10.764000 0.135590 0.397829 0.285534 3.464181 3.607500 19.953923 303 578.000000 408.692000 0.355431 0.479048 0.330865 4.863494 3.102771 19.198684 304 363.000000 906.071600 0.528777 0.597128 0.357216 6.490117 2.855443 19.025489 .

Figure 5: Example regional.txt file.

Files ending in connectivity.txt contain the regional connectivity data, in the form of a square symmetric matrix of regional density values, where each row and column corresponds to one region, and rows and columns are ordered by the region code, starting at 300 (No label) and ending at 318 (L cerebellum) (refer to Table 1).

0.0674386290 0.0330784566 0.0204603581 0.0538503291 ... 0.0330784566 0.0278032082 0.0129259694 0.0371359991 ... 0.0204603581 0.0129259694 0.0277173913 0.0171291023 ... 0.0538503291 0.0371359991 0.0171291023 0.0820195139

Figure 6: Example connectivity.txt file.

2.3 Connectivity analysis

The file connectivity_analysis.tbz2 contains regional connectivity data, and the results of statistical tests between each pair of groups upon regional connectivity. Once decompressed, the data is organised as follows:

```
connectivity_analysis/
concat/
trialavg/
```

Inside each of the concat and trialavg sub-directories are two types of plain text files, named as follows:

```
[space]_[gtype]_[ttype].txt
test_[space]_[gtype]_[ttype].txt
```

where:

- [space] is one of fmri or std.
- [gtype] is one of whole, nocsf, gm or wm.
- [ttype] is one of abs, pos or neg.

Files which have a name of the form [space]_[gtype]_[ttype].txt contain intra- and inter-regional connectivity values for every subject. An example is shown in Figure 7. Each row contains the connectivity data for every region and pair of regions, for a specific network class from one subject.

Files with a name of the form test_[space]_[gtype]_[ttype].txt contain the results of the statistical tests between each pair of groups upon the connectivity data. An example is shown in Figure 8. Each row contains the group means

#	groups:	young: 0, awout: 1, awith: 2	_
#	thresholds:	0.818398: 0, 0.899876: 1, 0.962249: 2	
#	group" "subject"	"threshold" "300" "301" "302" "300-301" "300-302" "300-303"	
0	16 0 0.07207040 0	.03211864 0.03548137 0.03919045 0.02844964 0.05889816	
0	16 1 0.02337502 0	.00781879 0.01011066 0.00968739 0.00734986 0.01500545	
0	16 2 0.00793057 0	.00290365 0.00765027 0.00242017 0.00249900 0.00253942	
.			
.			

Figure 7: Example connectivity data file.

and standard deviations, and the t value and two tailed probability for a test at a single region or pair of regions. As part of correction for multiple comparisons, any results which were not found to be significant were given a probability of 1.0.

Figure 8: Example connectivity test file.

2.4 Global analysis

The file global_analysis.tbz2 contains global network measure data, and the results of statistical tests upon all global network measures, between each pair of groups. Once decompressed, the data is organised as follows:

global_analysis/

Inside each of the concat and trialavg sub-directories are two types of plain text files, named as follows:

```
[space]_[measure].txt
test_[space]_[measure].txt
```

where:

- [space] is one of fmri or std.
- [measure] is the name of the relevant global network measure (refer to Table 7).

Files which have a name of the form [space]_[measure].txt contain global network data for the specified measure, for every subject. These files begin with a header section which describe the data, with each subsequent line giving the measure data for one subject and network class. An example is given in Figure 9.

```
# groups: young: 0, awout: 1, awith: 2
# graph types: gm: 0, nocsf: 1, whole: 2, wm: 3
# threshold types: abs: 0, neg: 1, pos: 2
# thresholds: 0.446347: 0, 0.534151: 1, 0.629034: 2
# "group" "subject" "graph type" "threshold type" "threshold" "pruned"
0 16 0 0 0 33
0 16 0 0 1 150
0 16 0 0 2 959
.
.
.
```

Figure 9: Example global measure data file.

Files with a name of the form test_[space]_[measure].txt contain the results of the statistical analysis upon the specified global network measure. In a similar manner to the global measure data files, these files begin with a header section, and then contain one line for each statistical test which was carried out. Each line contains the group means, standard deviations and number of samples (the grp1mean, grp1std and grp1n columns respectively, and the corresponding columns for group 2), and the U value and corresponding two-tailed probability from the Mann-Whitney U test. While not reported in Section 1 or in [12], a two sample t test was also carried out upon the data – the resulting t values and two tailed probabilities are given in the tval and tprob columns respectively.

```
# groups: young: 0, awout: 1, awith: 2
# graph types: gm: 0, nocsf: 1, whole: 2, wm: 3
# threshold types: abs: 0, neg: 1, pos: 2
# thresholds: 0.446347: 0, 0.534151: 1, 0.629034: 2
# "grp1" "grp2" "graph type" "threshold type" "threshold" "grp1mean" "grp1std" "grp1n" "grp2mean" "grp2std" "grp2n" "uval" "uprob" "tval" "tprob"
0 1 0 0 0 73 51 14 39 36 14 47 0.02157896 2.064438 0.04908556
0 1 0 0 1 365 233 14 215 164 14 51 0.03260942 1.964734 0.06021612
0 1 0 0 2 1310 704 14 969 423 14 67 0.16103591 1.555523 0.13191120
.
```

Figure 10: Example global measure test file.

2.5 Regional analysis

The file **regional_analysis.tbz2** contains the results of statistical tests upon all regional network measures between each pair of groups. Once decompressed, the data is organised as follows:

```
regional_analysis/
_____ concat/
_____ trialavg/
```

Inside each of the concat and trialavg sub-directories are two types of plain text files, named as follows:

```
[space]_[measure].txt
test_[space]_[measure].txt
```

where:

- [space] is one of fmri or std.
- [measure] is the name of the relevant regional network measure.

Files for measures to which Z-normalisation has been applied end in _z.txt. Files which have a name of the form [space]_[measure].txt contain regional network data for the specified measure, for every subject; an example file is shown in Figure 11. After a header section describing the file, each line contains the data for all regions from a one network class for one subject.

# groups:	young: 0, awout: 1, awith: 2
# graph types:	gm: 0, nocsf: 1, whole: 2, wm: 3
<pre># threshold types:</pre>	abs: 0, neg: 1, pos: 2
# thresholds:	0.818398: 0, 0.899876: 1, 0.962249: 2
# "group" "subject"	"graph type" "threshold type" "threshold" "300" "301" "302" "303"
0 16 0 0 0 0.133890	1.020150 2.036166 0.062535
0 16 0 0 1 0.089406	1.018155 2.162590 -0.027799
0 16 0 0 2 0.166253	0.978073 2.138208 0.247636

Figure 11: Example regional measure data file.

Files with a name of the form $test_[space]_[measure].txt$ contain the results of the statistical analysis upon the specified regional network measure. As shown in Figure 12, each line contains the group means, standard deviations and number of samples for the measure at a single region and the t value and two tailed probability resulting from the statistical test comparing the data from the two groups. As part of correction for multiple comparisons, any results which were not found to be significant were given a probability of 1.0.

Figure 12: Example regional measure test file.

2.6 Voxelwise analysis

The file voxelwise_analysis.tbz2, contains the results of every voxelwise statistical test that was conducted upon the functional network data, along with anatomical label and tissue classification image files for every subject. Once this file is decompressed, the data is organised as follows:



where

- group_pair is one of young_awout, young_awith or awout_awith.
- space is one of fmri or std.
- graph_type is one of whole, nocsf, gm or wm.
- threshold_type is one of abs, pos or neg.
- threshold is one of 0.818398, 0.899876 or 0.962249 for task based (trialavg) networks, or 0.446347, 0.534151 or 0.629034 for resting state (concat) networks.

ANALYZE75 image files containing anatomical labels for each subject are located in the labels/anatomy directory. Image files containing tissue labels (grey matter/white matter/CSF) are located in the labels/tissue directory. Image files containing results of the voxelwise statistical tests are located in the threshold subdirectories. All ANALYZE75 images adhere to the orientation, resolution and dimensions outlined in Table 2, and are stored in little endian format.

Inside each of the threshold subdirectories, 12 images are present⁷ for each voxelwise measure (recall Tables 5 and 7):

[measure]_t.img
[measure]_corp.img
[measure]_dof.img

 $^{^7\}mathrm{Every}$ file ending in .img is accompanied by a corresponding .hdr file.

[measure]_fill_t.img
[measure]_fill_corp.img
[measure]_z_t.img
[measure]_z_corp.img
[measure]_z_dof.img
[measure]_z_fill_t.img
[measure]_z_fill_corp.img

where:

- [measure] is the name of the voxelwise measure.
- Files ending in _t.img contain the two-sample t values at every voxel.
- Files ending in _corp.img contain the (corrected) two-tailed probabilities at each voxel.
- Files ending in _dof.img contain the degrees of freedom used in the *t*-test at each voxel.
- Files with measure names ending in $_{z}$ contain the test results for Z-normalised network measure data.
- Files with measure names ending in _fill contain the test results for imputed network measure data.
- Files with measure names ending in _z_fill contain the test results for Z-normalised, imputed network measure data.

Anatomical label images (recall Table 1) are located in the labels/anatomy directory. Two image files are present for each subject:

```
[group]_[subject_no]_fmri_labels.img
[group]_[subject_no]_fmri_labels_reduce.img
```

where [group] is one of young, awout or awith, and [subject_no] is the two digit subject identifier. Files ending in _labels.img contain labels corresponding to the original LPBA40 anatomical parcellation, and files ending in _labels_reduce.img contain labels corresponding to the reduced parcellation used in our analysis. These label images are in the native space of each subject, and thus correspond to native/fMRI space functional networks. Anatomical label images are also provided for standard space networks:

```
std_labels.img
std_labels_reduce.img
```

These images are applicable to standard space functional networks for all subjects.

The labels/tissue directory contains tissue classification images for each subject, in both native and standard space. Two images are present for each subject:

```
[group]_[subject_no]_fmri_tissue.img
[group]_[subject_no]_std_tissue.img
```

where [group] is one of young, awout or awith, and [subject_no] is the two digit subject identifier. Inside these images, a value of 1 corresponds to cerebrospinal fluid, 2 to grey matter, and 3 to white matter. A value of 0 corresponds to background. As described in Section 1.2, these tissue classification images were used as binary masks in the creation of whole brain, white+grey matter, grey matter and white matter functional networks.

References

- Andersson, J. L., Jenkinson, M., and Smith, S. (2007). Non-linear registration, aka Spatial normalisation. FMRIB Technical Report TR07JA2, FMRIB Centre, Oxford, United Kingdom.
- [2] Buckner, R. L., Sepulcre, J., Talukdar, T., Krienen, F. M., Liu, H., Hedden, T., Andrews-Hanna, J. R., Sperling, R. A., and Johnson, K. A. (2009). Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. J. Neurosci., 29(6), 1860–73.
- [3] Buckner, R. L., Snyder, A. Z., Sanders, A. L., Raichle, M. E., and Morris, J. C. (2000). Functional Brain Imaging of Young, Nondemented, and Demented Older Adults. *Journal of Cognitive Neuroscience*, 12(Supplement 2), 24–34.
- [4] Cox, R. W. (2012). AFNI: What a long strange trip it's been. Neuroimage, 62(2), 743–747.
- [5] fMRIDC (2011). The fMRI Data Center. http://www.fmridc.org.
- [6] Greicius, M. D., Srivastava, G., Reiss, A. L., and Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: Evidence from functional MRI. PNAS, 101(13), 4637–4642.
- [7] Hayasaka, S. and Laurienti, P. J. (2010). Comparison of characteristics between region-and voxel-based network analyses in resting-state fMRI data. *NeuroImage*, 50(2), 499–508.
- [8] Hayasaka, S. and Nichols, T. E. (2003). Validating cluster size inference: random field and permutation methods. *NeuroImage*, 20(4), 2343–2356.
- [9] Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., and Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry*, 140(6), 566–572.
- [10] Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2), 825–841.
- [11] Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Le Goualher, G., Boomsma, D., Cannon, T., Kawashima, R., and Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, 356(1412), 1293–322.
- [12] McCarthy, P. (2013). Functional network analysis of Aging and Alzheimer's Disease. Ph. D. thesis, University of Otago, Dunedin, New Zealand.
- [13] Nichols, T. E. and Holmes, A. P. (2001). Nonparametric Permutation Tests for Functional Neuroimaging: A Primer with Examples. *Human Brain Mapping*, 15(1), 1–25.
- [14] Shattuck, D. W. and Leahy, R. M. (2002). BrainSuite: An automated cortical surface identification tool. Medical Image Analysis, 6(2), 129–142.
- [15] Shattuck, D. W., Sandor-Leahy, S. R., Schaper, K. A., Rottenberg, D. A., and Leahy, R. M. (2001). Magnetic Resonance Image Tissue Classification Using a Partial Volume Model. *NeuroImage*, 13(5), 856–876.
- [16] Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., 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., and Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23 Suppl 1, S208–19.
- [17] Snyder, A. (2011). private communication.
- [18] Vaden Jr, K. I., Gebregziabher, M., Kuchinsky, S. E., and Eckert, M. A. (2012). Multiple imputation of missing fMRI data in whole brain analysis. *NeuroImager*, 60(3), 1843–1855.
- [19] van den Heuvel, M. P., Stam, C. J., Boersma, M., and Hulshoff Pol, H. E. (2008). Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. *Neuroimage*, 43(3), 528–539.