Neurovascular imaging markers of brain aging

Sigurdsson, S.

2

The AGES-Reykjavik study atlases: Non-linear multi-spectral template and atlases for studies of the ageing brain

Manuscript based on this chapter has been published as:

Forsberg L, Sigurdsson S, Fredriksson J, Egilsdottir A, Oskarsdottir B, Kjartansson O, van Buchem MA, Launer LJ, Gudnason V, Zijdenbos A. The AGES-Reykjavik study atlases: Non-linear multi-spectral template and atlases for studies of the ageing brain. Med Image Anal. 2017 May 6;39:133-144.

ABSTRACT

Quantitative analyses of brain structures from Magnetic Resonance (MR) image data are often performed using automatic segmentation algorithms. A common approach to achieve brain structure segmentation is to estimate a spatial transformation between a model brain ("template") and an individual subject, and subsequently use that transformation to map an atlas defined on the template to the individual brain. Most freely available brain atlases are generated from relatively young individuals and not always derived from well-defined cohort studies. In this paper, we introduce a publicly available template and a whole brain atlas with cortical and subcortical regions together with a regional segmentation pipeline, which are optimised to use in studies of ageing cohorts (mean age 76±6 years). Furthermore, we provide validation data to assure accuracy and reproducibility of the segmentation results.

INTRODUCTION

Over the last two decades, a number of methods have been introduced to map the human brain. Many of these use atlas based techniques to analyse the brain functionally and structurally.[1-5] Average intensity atlases describe the average signal intensity in a common coordinate space and are often constructed from T1-weighted Magnetic Resonance Image (MRI) scans and referred to as standard "templates". They can either be constructed using a linear or a non-linear transformation of the individual subjects to the common space, where the images are averaged. Many of these are symmetrical, meaning the left and the right hemispheres are forced to be mirror images. The most commonly used of these is the ICBM152 standard template, typically used as a registration target in functional and structural group studies. This template was constructed using 152 brain scans acquired at the Montreal Neurological Institute for the International Consortium for Brain Mapping (ICBM) project. It is the successor to the older MNI305 template, which was built by averaging 305 linearly registered T1-weighted MR scans. The ICBM152 template is available in both linear and nonlinear, symmetric and asymmetric versions, and includes T1-, T2-, PD-weighted intensity atlases and tissue probability atlases.[1,6-7] Tissue probability atlases describe the likelihood that a certain voxel in a template space belongs to a specific tissue. These tissue probability atlases can be used as priors in the tissue segmentation of grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) in individual subjects.[8] Regional probability atlases divide the brain into a number of brain regions and can describe which region is most likely for each voxel in the brain. These can be used for regional segmentation of individual subjects or as a reference atlas in template space. Regional segmentation can be achieved by warping regions of interest from atlas space to the individual subjects, possibly while further improving the segmentation by taking into account the classified tissues in subject space.[9] One commonly used regional atlas is the AAL-atlas[10], which is part of the IBASPM toolbox for the SPM software package. This atlas is based on the MNI single-subject template and consists of 90 anatomical regions. Another single-subject based atlas is the MNI structural atlas. In this atlas, the labels are non-linearly registered to the structural images of more than 50 subjects and then transformed to ICBM152 space to finally produce the probability images, thereby taking into account the morphometric variability across subjects.[11] This atlas is part of the collection of atlases that come with the FMRIB Software Library (FSL).[12] An alternative atlas is the LONI Probabilistic Brain Atlas (LPBA40), which consists of 56 regions based on scans of 40 subjects with an average age of 29.2 years.[13] In this atlas, the structures where manually labelled in each subject after registration to the MNI305 template to form the probabilistic atlas.

A number of studies have identified a need for population-specific brain templates and atlases. For instance, Machilsen et al.[14] showed that the ICBM152 template is not ideal for pediatric studies since it may introduce inaccuracies or bias in the spatial normalisation. This problem was addressed by creating unbiased age-appropriate pediatric tissue probability atlases.[7] Similarly for ageing studies, Mega et al.[15] have created a probabilistic brain atlas from an elderly cohort with dementia, which is better suited for studying Alzheimer's disease. Also, in Voxel Based Morphometry studies, it is common practise to create study-specific templates to avoid registration bias.[16] In ageing studies, another potential bias is the misclassification of brain tissue due to white matter lesions. These lesions appear hypointense in the T1-weighted image, which may lead to overestimation of grey matter in white matter regions when only relying on T1-weighed images.[17] Bias may therefore be introduced in both tissue atlases and regional atlases when using automatic tissue segmentation to delineate between grey matter and white matter regions.

In this study, we have created an ageing-specific T1-weighted template and an anatomical atlas with 56 regions based on 314 subjects from the AGES-Reykjavik study (mean age 76±6 years).[18] These are called the AGES-template and the AGES-atlas. We addressed the problem of hypo-intense white matter lesions by taking these into account as a fourth tissue class in the generation of the atlas. A fully automatic regional segmentation pipeline, which uses the atlas, was also created. Finally, we validated the template, atlas and segmentation pipeline to assure accuracy and reproducibility of the results. This regional atlas can also be used as a reference in template space.

METHODS

Study sample

The AGES-Reykjavik Study cohort consists of 5764 participants, 4811 of which underwent brain MRI. All MRI scans were processed using the tissue segmentation pipeline described by Sigurdsson et al.[19] Of these, 4614 scans passed quality control of the automatic brain tissue segmentation (1934 men, 2680 women, mean age 76 ± 6 years). For the purpose of constructing a population average, 400 subjects were initially randomly selected from the 4811 subjects, out of which 86 subjects were subsequently removed due to large brain infarcts or poor quality in the tissue segmentation, leaving a pool of 314 scans (124 men, 190 women, mean age 75 ± 5 years, age range 66 to 92 years). Being part of the population, the decision was to include cases with dementia and MCI (Mild Cognitive Impaired) if they were selected by the random selection process. The 314 scans were used to construct an anatomical minimum-deformation template[7] and

to generate probabilistic atlases. Another group of 31 subjects from the cohort of 4614 subjects with valid scans were randomly selected for reproducibility experiments (18 men, 13 women, mean age 75 ± 5). This group underwent a same-day repeat scan, where the subjects were removed from the scanner between scans. These data were used for reproducibility testing. Finally, another 7 scans were randomly selected from the 4614 and manually segmented into 56 anatomical regions-of-interest (see Table 3 for the list of regions), where 4 subjects (2 men, 2 women, mean age 74) were used to construct the initial atlas and 3 subjects (1 man, 2 women, mean age 80) were used for accuracy testing. Subjects that were used for reproducibility testing or accuracy testing were not part of the 314 subjects used to generate the template and the regional atlas. All MR images were acquired using a dedicated General Electrics 1.5-Tesla Signa Twinspeed EXCITE system (Waukesha, WI) with a multi-channel phased array head cap coil, using the following image parameters: T1-weighted (TE, 8 ms; TR, 21 ms; FA, 30; FOV, 240 mm; matrix, 256x256) with 1.5 mm slice thickness and 0.94 mm x 0.94 mm in-plane pixel size, proton density (PD)/T2-weighted fast spin echo (FSE) sequence (TE1, 22 ms; T2, 90 ms; TR, 3220 ms; echo train length, 8; FA, 90; FOV, 220 mm; matrix, 256x256), and fluid attenuated inversion recovery (FLAIR) sequence (TE, 100 ms; TR, 8000 ms; inversion time, 2000 ms, FA, 9; FOV, 220 mm; matrix, 256x256) were acquired. Proton density (PD)/T2-weighted and FLAIR were acquired with 3 mm slice thickness and 0.86 mm x 0.86 mm in-plane pixel size.

Image pre-processing and tissue segmentation

The MR images were processed through a tissue segmentation pipeline, described in detail by Sigurdsson et al.[19] In brief, the T1-weighted, PD/T2-weighted and FLAIR images were first each corrected for signal non-uniformity using the N3 algorithm.[20] The PD/T2-weighted and FLAIR images were then co-registered to the T1-weighted image and the image set was subsequently spatially normalised to the MNI-ICBM152 template[3] using a multi-resolution optimization algorithm yielding a 9-parameter ane transformation.[21] Using this affine spatial transform, the images were resampled to a 1 mm isotropic resolution, linearly intensity-normalised, and used as input to a trained artificial neural network tissue classier to segment the brains into cerebrospinal fluid (CSF), grey matter (GM), normal white matter (NMW), and white matter hyperintensities (WMH).[22] White matter (WM) was regarded as NMW+WMH. Skull removal was done using BET (Brain Extraction Tool).[23] The final tissue segmentation and skull removal results were quality controlled through visual inspection of each subject. In the following, the term “MNI space” will be used to denote the 1mm isotropic sampling lattice of the spatially normalised images.

Template generation

To obtain an unbiased registration- and segmentation template image representative of the AGES population, a minimum-deformation T1-weighted template was generated from the 314 intensity normalised, T1-weighted images in MNI space, using the methodology described by Fonov et al.[7] In brief, this method cycles through registration- and averaging phases; in the registration phase, each image is registered to the voxel-wise image average generated in the previous cycle. In order to remove any left-right differences in the template, each image average was explicitly symmetrised by averaging it with a left-right flipped version of itself. The procedure was initialised by first creating a voxel-wise average from the MNI space spatially normalised T1-weighted images of the 314 template subjects. Following the initial linear registration phase, subsequent registration phases were performed using non-linear registration through the estimation of a deformation field. As is commonly done in non-linear registration, the source- and target images were blurred using a Gaussian kernel with a full width at half maximum (FWHM) ranging from 8 mm to 1 mm, allowing for a multi-resolution approach to the spatial deformation through the registration/averaging cycles.[24] To improve convergence, the registration/averaging cycle was repeated twice for the early registration phases and four times for the later registration phases as shown in Table 1.

Table 1. The registration/averaging cycles used in the template generation process, showing the number of times a registration/averaging cycle was run at each level of spatial normalization (blurring). See also Figure 5.

	FWHM (mm)					
	Linear	8	6	4	2	1
#	1	2	2	4	4	4

In order to assess convergence, the standard deviation volume was calculated at each iteration. The standard deviation is expected to decrease across iterations, as the individual deformed images align better. The iterative procedure was monitored to ensure that the standard deviation monotonically decreased and thus that the process converged. The resulting minimum-deformation template is shown in Figure 5, and defines the “AGES space.”

The final subject-to-template deformations were also used to warp and construct corresponding T2-weighted, PD-weighted and FLAIR templates as well as tissue probability maps. The templates were constructed by warping the images to AGES space where they were symmetrised and averaged. The same procedure was used to warp each subject’s CSF, WM, WM, and WMH segmentations to AGES space and construct corresponding global tissue probability maps.

Regional atlas generation

To generate an anatomical atlas, 56 anatomical regions-of-interest (ROIs; see Table 3) were manually labelled on the MRI scans of 4 subjects by an anatomical expert using an inhouse developed labelling tool. In order to make the labelling procedure efficient, the expert was able to use the tissue maps from the automatic tissue segmentation as a guideline and could choose to follow the borders from the tissue segmentation or to draw freehand. Subcortical regions, like Thalamus and Putamen, were often drawn freehand while cortical regions were drawn using the guidelines. It took about a week to manually label the whole brain for each subject.

The T1 image for each of the 4 subjects was non-linearly warped to the AGES template and the resulting deformation fields were subsequently used to warp the manually labelled regions into AGES space. Each region was separately warped using trilinear interpolation, initially resulting in "fuzzy" ROI volumes. The individual regions were then left-right mirrored and averaged with the non-mirrored to construct a symmetric initial atlas, where each voxel describes which region that is most likely. MRI scans for 3 other subjects were manually labelled for validation purposes and were not used to create the atlas.

The initial atlas was used for regional segmentation of the template subjects. This was done by first warping the template to each subject's T1-weighted image and then applying the same deformation field to the atlas in order to warp the atlas containing all regions to each individual subject using nearest neighbour interpolation. The regions were then constrained by the tissue segmentation of the individual subjects so that GM voxels would not get classified as being WM or CSF regions. A set of rules allowed GM voxels classified as a WM regions to be reclassified as the corresponding GM regions (and vice versa for WM voxels). Voxels where no reclassification could be made were left unlabelled. The segmented image containing all regions was warped back to the AGES space using nearest neighbour interpolation. In AGES space, each region was mirrored to construct a symmetric population-based probability atlas based on the automatic segmentation of the 314 template subjects. Regional probability atlases were finally created for CSF, GM and WM, where the most probable region is stored for each voxel.

Regional segmentation pipeline

The regional segmentation pipeline is shown in Figure 1. It starts with warping the AGES template to the subject's T1-weighted image, where the deformation field obtained is used to warp the regional CSF, GM and WM atlases to the subject's T1-weighted image. The 3 warped regional tissue atlases are then multiplied with the subject's corresponding tissue masks (previously obtained from the global tissue segmentation pipeline),

and the results are summed together to form the final regional segmented image. This procedure is similar to the ANIMAL+INSECT method by Collins et al.[9]

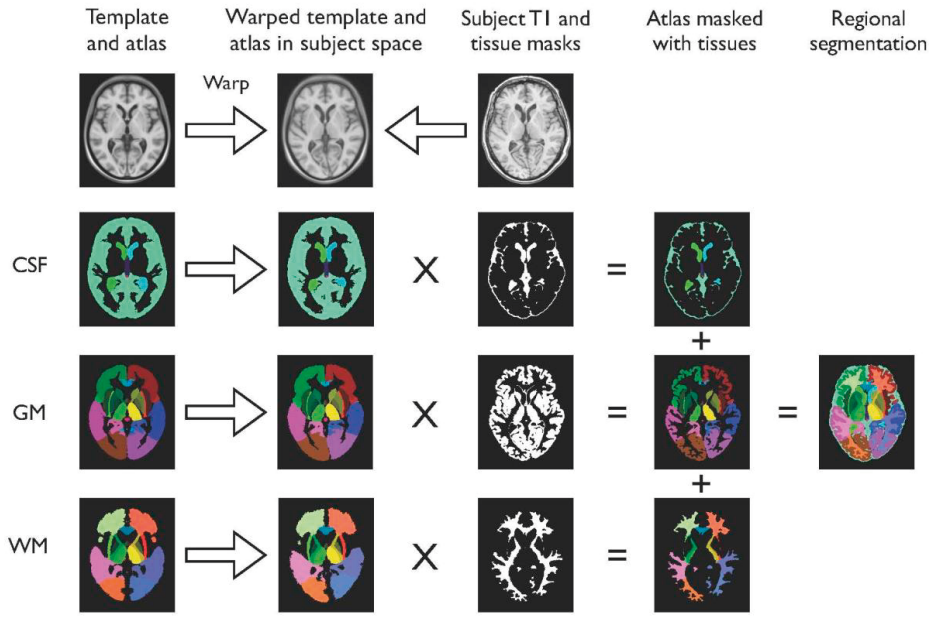

Figure 1. Regional segmentation pipeline. The AGES template is warped to the subject's T1-weighted image and the deformation field is used to warp the corresponding tissue atlases, which are then multiplied by the subject's tissue masks. The results are finally added together to form the final regional segmentation image.

Regional atlas and pipeline validation

The validation was divided into two parts: reproducibility and accuracy. Reproducibility measurements were obtained using the repeated scans of the 31 subjects, where each subject was scanned twice the same day and the two visits were processed through the regional segmentation pipeline independently. The repeated scans were used for testing the reproducibility of the whole process, from the generation of the images in acquisition to the results from the regional segmentation pipeline. The Dice Similarity Coefficient (DSC; also known under other names, such as the Sörensen Index) was calculated for each subject and region to test the agreement of the repeated scans on a subject level, where DSC>0.70 is generally considered excellent agreement.[22] DSC is given by:

$$\mathrm{DSC} = 2\frac{\mathrm{A} \cap \mathrm{B}}{\mathrm{A} \cup \mathrm{B}} \quad \textbf{(1)}$$

where A and B are the regional volumes to be compared, in this case between the first and the second visit for a given region. This measurement gives a value between 0 and 1 for each region. A value close to 1 indicates a strong agreement.

The within- and between-subject Coeffcient of Variations (CV), a commonly used parameter of measurement variability,[25,26] were calculated to assess regional volume reproducibility. In contrast to standard deviation, which must always be understood in relation to the mean, the CV metric is a normalised variance in percentage for each region. This makes it possible to compare the CV values between regions. The within-subject CV is given by

$$CV_w = 100\frac{\sigma_w}{\mu\sqrt{2}} \quad (2)$$

where σ_W is the standard deviation of the difference between the two visits across subjects and μ is the average volume of the region for both visits across all subjects.[27] The between-subject CV is given by

$$CV_B = 100\frac{\sigma_B}{\mu} \quad (3)$$

where σ_B is the standard deviation of the average between the two visits across subjects and μ is the average volume of the region across all subjects. Since the subjects are scanned the same day, we should expect the within-subject CV_W to be lower than the between-subject CVB. To assess and compare the CV performance across regions, the Index of Individuality ratio IOI = CV_W/CV_B was calculated. For reproducibility testing, where there should be no differences between visits, we expect to get a low IOI defined as IOI < 0.60.[28] The Spearman correlation between mean reproducibility DSC and regional volume across regions was finally calculated to see if reproducibility results are driven by volume. This was done both with and without Pineal Gland, which is a very small region compared to all other regions and could therefore be considered an outlier volume wise.

Accuracy measurements from manually segmented images were obtained using brain images of 3 subjects that were labelled by an anatomical expert. The accuracy was tested by calculating DSC for each region between the manual and automatic segmentations of these subjects, and the Spearman correlation between mean accuracy DSC and the regional volume across regions was calculated to see if the accuracy results are driven by volume, again with and without Pineal Gland.

RESULTS

Template and atlases

The final template and corresponding probability atlases are shown in Figure 2. Table 2 provides the characteristics of the 314 subjects. The four rows in Figure 2 show the T1-weighted template, regional CSF probability atlas, regional GM probability atlas, regional WM probability atlas, and the final regional atlas that includes all regions. The regional tissue probability atlases are used in the regional segmentation pipeline, as shown in Figure 1, and capture the inter-subject variability in AGES space, therefore overlapping each other. The final atlas is not used by the tissue segmentation pipeline but can serve as a general purpose lookup atlas in linear MNI space or in non-linear AGES space. Figures 3 and 4 show the T2-weighted, PD-weighted and FLAIR templates, and the global tissue probability maps.

Table 2. Characteristics of the template subjects (n=314) by sex, with standard deviation and range.

Demographics	Overall, n = 314	Men, n = 124	Women, n = 190
Age	76.2 ± 5.4 [66,92]	76.4 ± 5.2 [67,88]	76.1 ± 5.6 [66,92]
Body Mass Index	26.9 ± 4.2 [16.9,44.8]	26.7 ± 3.8 [18.2,37.5]	26.9 ± 4.5 [16.9,44.8]
Cognitive status			
Normal	265 (84.4%)	103 (83.1%)	162 (85.7%)
Mild Cognitive Impaired	33 (10.5%)	10 (8.1%)	23 (12.2%)
Dementia	15 (4.8%)	11 (8.9%)	4 (2.1%)
Hypertension status			
No	13 (4.1%)	5 (4.0%)	8 (4.2%)
Previous	52 (16.6%)	16 (12.9%)	36 (18.9%)
Current	249 (79.3%)	103 (83.1%)	146 (76.8%)

The T1-weighted template was generated through 17 cycles, where cycle 17 corresponds to the final template. Figure 5 shows the progression of the average template and the standard deviation at the end of each blurring level, and Figure 6 shows the standard deviation at each cycle, calculated as the square root of the average variance across all brain voxels. The anatomical features got sharper for each cycle while the standard deviation across subject decreased for the whole brain as the template became sharper. Both figures show that the template converged.

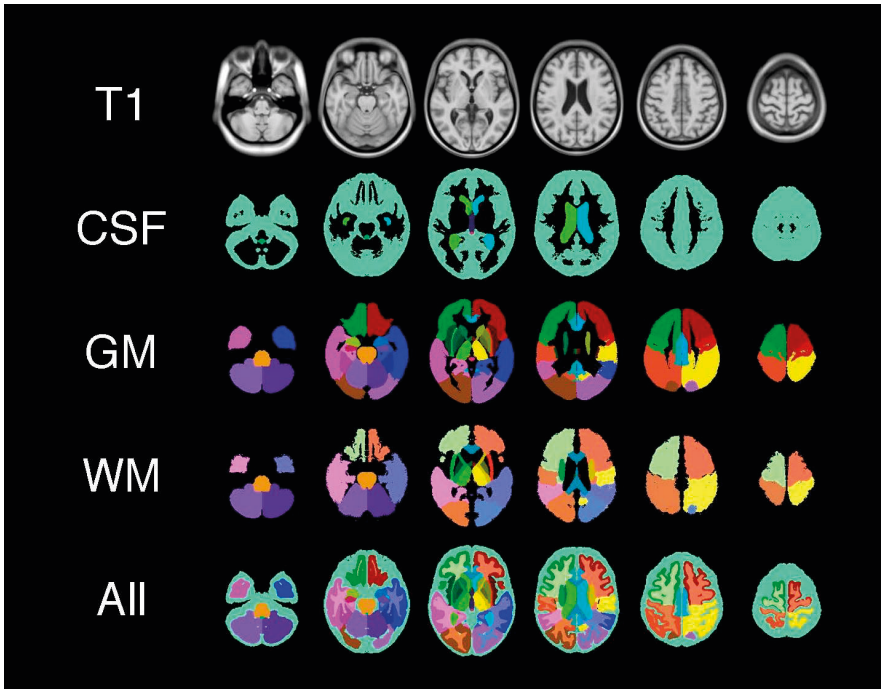

Figure 2. From top to bottom: The final template (T1-weighted), CSF regional tissue atlas, GM regional tissue atlas, WM regional tissue atlas, and finally the combined regional atlas for all tissues. The combined atlas is not used by the regional segmentation pipeline but can be used as a lookup-atlas in linear MNI space or non-linear AGES space.

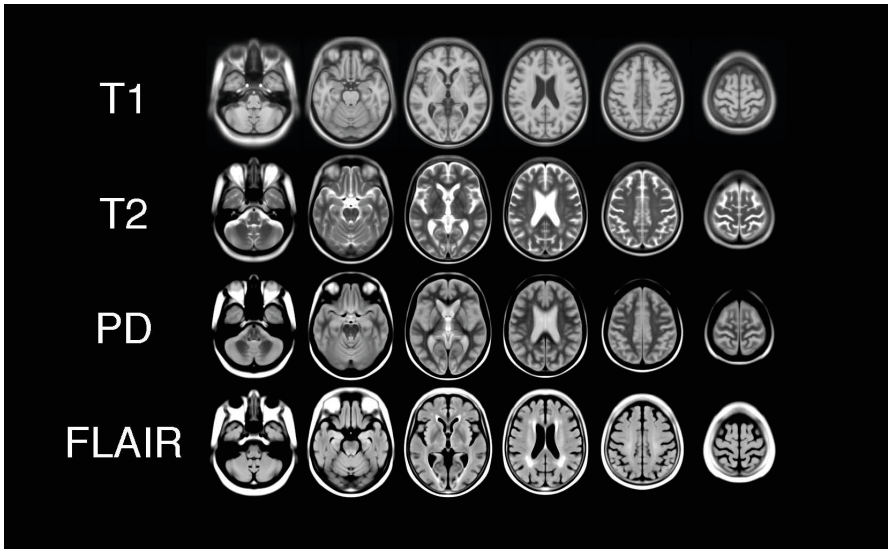

Figure 3. The templates for the four image modalities: T1-weighted, T2-weighted, PD-weighted, and FLAIR.

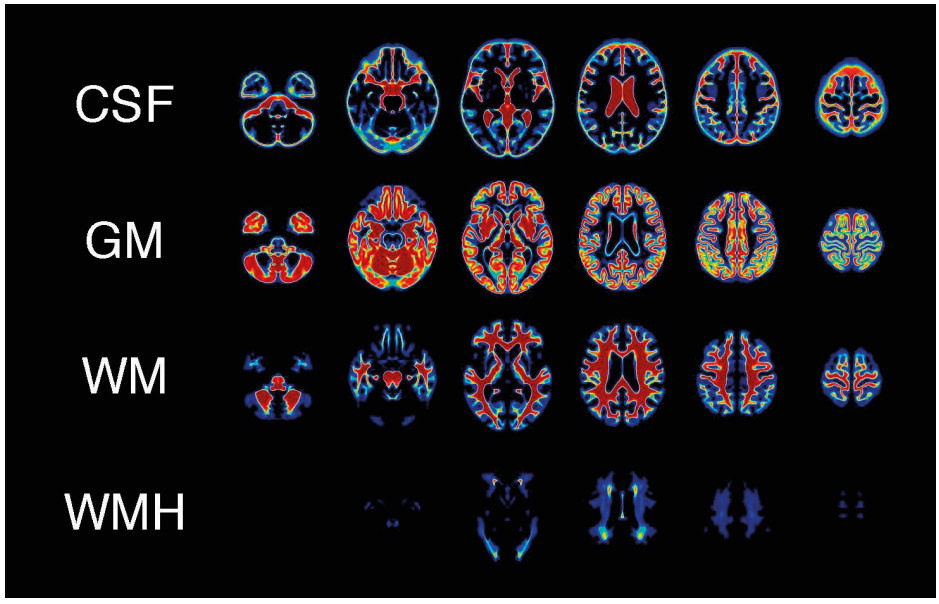

Figure 4. The global tissue probability maps: CSF, GM, WM and WMH.

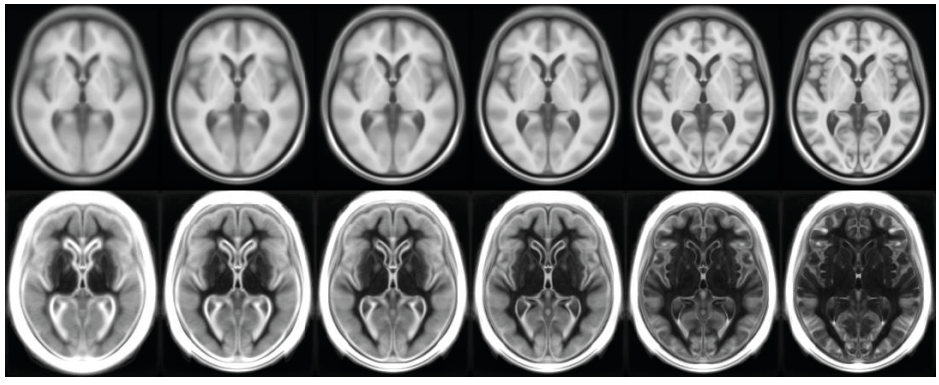

Figure 5. The progression of the average template (top row) and the standard deviation of the template (bottom row) at the end of each blurring level. From left to right: Initial linear registration phase, 8 mm FWHM, 6 mm FWHM, 4 mm FWHM, 2 mm FWHM, and 1 mm FWHM blurring level.

Reproducibility results

The reproducibility results are shown in Table 3 and in Figures 7-9. Pineal Gland is included in the table but excluded from the figures to make them comparable with the accuracy figures where Pineal Gland is excluded due to its low DSC value. The figures show the DSC values for the 31 subjects in box plots. The median DSC value was higher than the mean value for all regions and a majority of regions (50 out of 56) had a DSC value above 0.8 while the remaining 6 regions had a DSC above 0.7. The IOI ratio showed that all regions

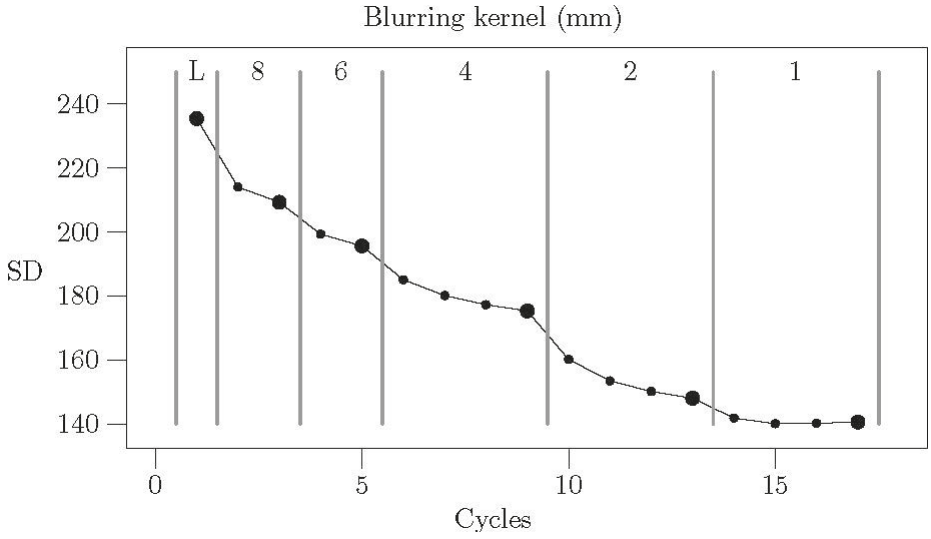

Figure 6. The template's standard deviation for each cycle, calculated as the square root of the average variance across all brain voxels. The larger dots represent the end of each blurring level. The first dot represents the initial linear registration (L).

Table 3. Reproducibility results for 31 subjects with two visits. Volume is calculated as the average volume across all subjects and visits.

Region	Avg DSC	Min DSC	Q1 DSC	Median DSC	Q3 DSC	Max DSC	CV_W (%)	CV_B (%)	IoI	Vol (mm^3)
L Orbitofrontal GM	0.838	0.571	0.810	0.860	0.882	0.931	1.843	9.526	0.193	74303
R Orbitofrontal GM	0.841	0.543	0.809	0.866	0.892	0.922	1.317	10.118	0.130	75182
L Precentral Gyrus	0.761	0.500	0.729	0.792	0.823	0.912	5.766	14.193	0.406	12962
R Precentral Gyrus	0.756	0.338	0.710	0.787	0.824	0.880	2.862	12.286	0.233	12618
L Lateral Temporal GM	0.859	0.720	0.826	0.878	0.901	0.927	1.314	11.380	0.116	47332
R Lateral Temporal GM	0.854	0.638	0.836	0.865	0.898	0.930	1.369	10.807	0.127	47343
L Medial Temporal GM	0.795	0.601	0.765	0.820	0.857	0.894	2.370	12.259	0.193	10731
R Medial Temporal GM	0.791	0.552	0.765	0.814	0.844	0.899	1.934	11.843	0.163	10834
L Lateral Occipitotemporal GM	0.832	0.701	0.804	0.845	0.883	0.919	2.440	11.082	0.220	5814
R Lateral Occipitotemporal GM	0.819	0.564	0.799	0.835	0.872	0.914	2.140	12.370	0.173	5811
L Occipital GM	0.832	0.688	0.785	0.850	0.882	0.920	2.365	14.004	0.169	42635
R Occipital GM	0.833	0.656	0.789	0.840	0.877	0.923	1.597	12.501	0.128	42950
L Parietal GM	0.802	0.574	0.768	0.834	0.858	0.929	2.368	12.992	0.182	42288
R Parietal GM	0.789	0.458	0.754	0.804	0.858	0.897	2.484	12.705	0.196	41139
L Insula	0.887	0.769	0.865	0.900	0.918	0.935	1.620	11.973	0.135	7068
R Insula	0.886	0.689	0.862	0.902	0.930	0.951	1.581	12.186	0.130	7023
L Cingulate	0.860	0.602	0.830	0.884	0.908	0.924	2.582	11.501	0.225	12376
R Cingulate	0.852	0.624	0.796	0.879	0.907	0.920	3.953	14.477	0.273	12671
L Caudate Nucleus	0.888	0.751	0.868	0.896	0.922	0.944	2.007	9.946	0.202	3601

Table 3. Reproducibility results for 31 subjects with two visits. Volume is calculated as the average volume across all subjects and visits. (continued)

Region	Avg DSC	Min DSC	Q1 DSC	Median DSC	Q3 DSC	Max DSC	CV_W (%)	CV_B (%)	IoI	Vol (mm^3)
R Caudate Nucleus	0.875	0.644	0.850	0.903	0.929	0.946	2.119	11.228	0.189	3508
L Putamen	0.902	0.813	0.893	0.909	0.931	0.950	2.077	10.767	0.193	4893
R Putamen	0.902	0.768	0.881	0.914	0.930	0.956	2.008	10.286	0.195	4819
L Globus Pallidus	0.844	0.659	0.816	0.871	0.894	0.935	4.564	12.334	0.370	1119
R Globus Pallidus	0.855	0.669	0.823	0.863	0.903	0.941	3.022	14.131	0.214	1108
L Thalamus	0.938	0.858	0.925	0.951	0.958	0.974	1.205	7.569	0.159	7548
R Thalamus	0.938	0.838	0.917	0.951	0.962	0.973	0.898	8.744	0.103	7660
L Accumbens	0.860	0.634	0.831	0.873	0.917	0.942	4.521	12.692	0.356	567
R Accumbens	0.861	0.663	0.808	0.868	0.934	0.959	3.374	14.467	0.233	556
L Amygdala	0.880	0.711	0.836	0.901	0.923	0.961	3.329	15.022	0.222	2336
R Amygdala	0.882	0.618	0.868	0.895	0.920	0.965	2.241	13.140	0.171	2497
L Hippocampus	0.867	0.717	0.833	0.897	0.918	0.937	1.464	13.415	0.109	2831
R Hippocampus	0.868	0.641	0.839	0.895	0.912	0.954	1.378	10.427	0.132	2768
Brain Stem	0.947	0.872	0.944	0.950	0.965	0.981	0.650	9.771	0.067	27746
L Cerebellum	0.948	0.897	0.937	0.953	0.964	0.976	0.741	11.994	0.062	60294
R Cerebellum	0.948	0.875	0.940	0.954	0.962	0.976	0.720	12.073	0.060	58801
L Frontal WM	0.885	0.745	0.865	0.898	0.916	0.953	1.753	12.301	0.143	69228
R Frontal WM	0.886	0.716	0.865	0.902	0.919	0.942	1.790	12.196	0.147	69226
L Temporal WM	0.849	0.714	0.818	0.866	0.887	0.924	2.315	14.222	0.163	31418
R Temporal WM	0.843	0.643	0.825	0.856	0.888	0.912	2.204	13.957	0.158	31513
L Occpital WM	0.860	0.755	0.829	0.877	0.897	0.928	2.912	14.318	0.203	26320
R Occipital WM	0.857	0.721	0.833	0.867	0.898	0.934	2.252	13.645	0.165	26534
L Parietal WM	0.863	0.714	0.839	0.881	0.903	0.949	5.056	14.507	0.348	36044
R Parietal WM	0.858	0.663	0.833	0.873	0.904	0.929	2.651	17.064	0.155	35276
L External Capsule	0.810	0.661	0.778	0.827	0.847	0.907	2.812	14.634	0.192	2490
R External Capsule	0.812	0.668	0.771	0.832	0.857	0.912	2.637	13.879	0.190	2428
L Int Caps	0.881	0.790	0.861	0.890	0.910	0.928	2.967	13.537	0.219	7509
R Int Caps	0.882	0.758	0.857	0.894	0.916	0.936	2.230	16.708	0.133	7724
L Fornix	0.810	0.563	0.747	0.856	0.882	0.914	3.709	14.274	0.260	822
R Fornix	0.808	0.577	0.767	0.848	0.881	0.928	3.248	17.442	0.186	652
Corpus Callosum	0.869	0.600	0.836	0.906	0.927	0.937	2.112	19.126	0.110	14267
L Lateral Ventricle CSF	0.945	0.899	0.937	0.946	0.961	0.972	1.348	55.117	0.024	22717
R Lateral Ventricle CSF	0.942	0.884	0.925	0.947	0.961	0.977	1.029	57.301	0.018	21518
Third Ventricle CSF	0.918	0.789	0.901	0.932	0.941	0.958	1.631	34.383	0.047	3489
Fourth Ventricle CSF	0.857	0.703	0.826	0.861	0.889	0.943	3.281	43.151	0.076	1677
Cortical CSF	0.826	0.675	0.807	0.842	0.869	0.888	4.618	27.017	0.171	387726
Pineal Gland	0.763	0.347	0.692	0.791	0.877	0.943	7.694	20.614	0.373	92

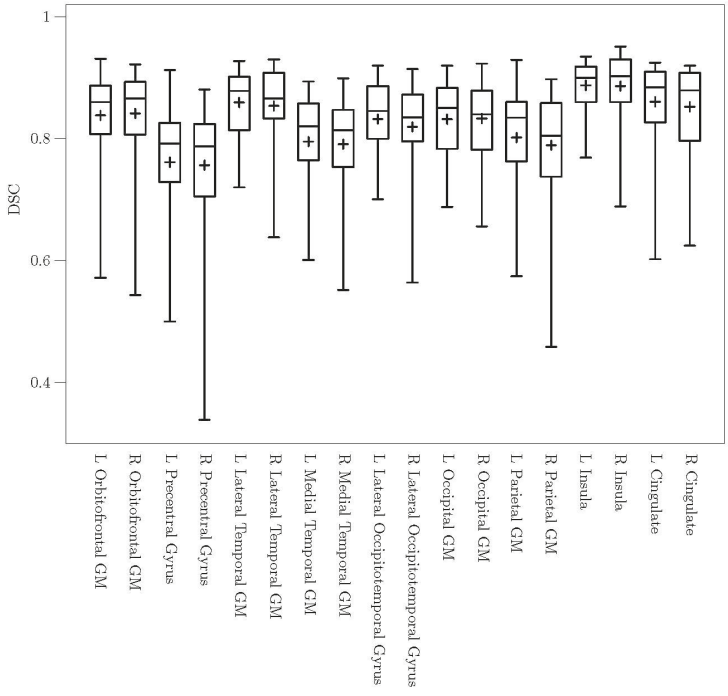

Figure 7. DSC reproducibility results for cortical regions.

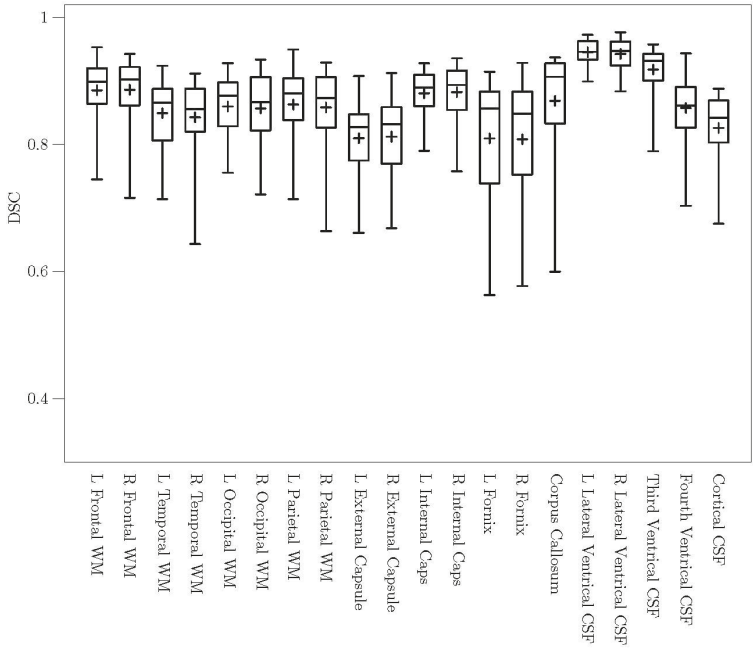

Figure 8. DSC reproducibility results for WM and CSF regions.

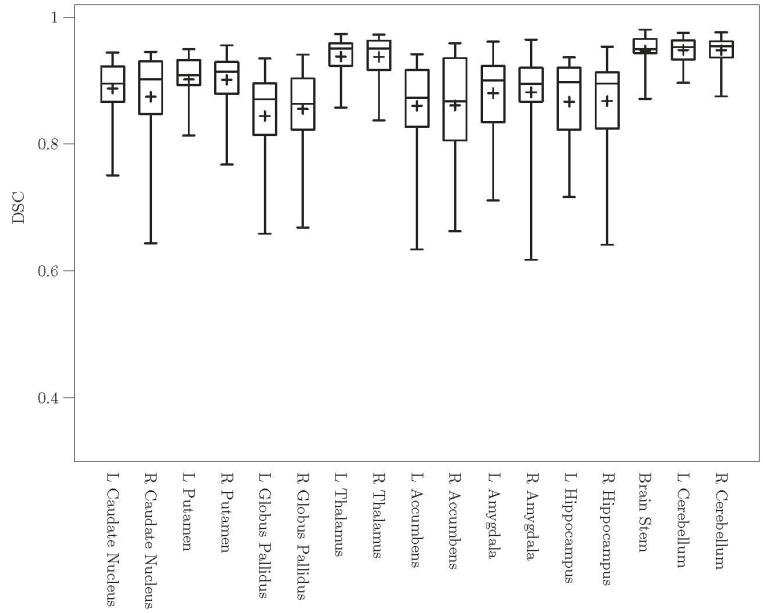

Figure 9. DSC reproducibility results for subcortical regions.

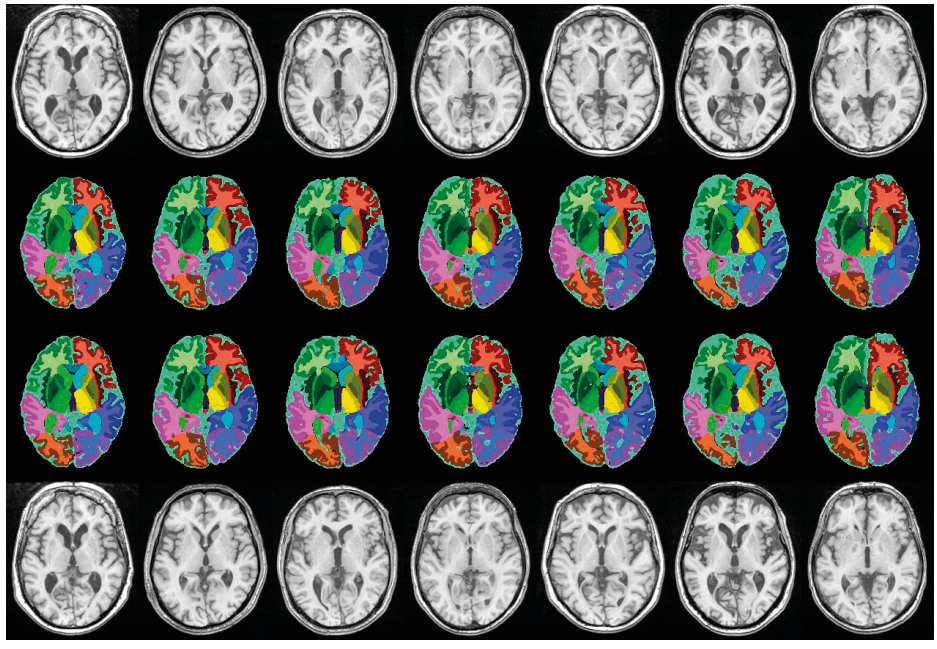

Figure 10. Repeated scans for seven subjects. The two first rows represent visit 1, the two last rows represent visit 2 from the same day.

had an IOI lower than 0.50. The Spearman correlation between mean DSC and volume was 0.042 with Pineal Gland and -0.006 without Pineal Gland, thus providing no evidence that reproducibility results should be driven by volume. Figure 10 demonstrates the performance of the regional segmentation for seven randomly selected repeated subjects. The two first rows represent visit 1 and the two last rows represent visit 2.

Accuracy results

The DSC results from the accuracy validation are shown in Table 4 and in Figures 11-13. The DSC range of the figures is from 0.5 to 1.0, which is why Pineal Gland with an average DSC of 0.227 is excluded in the figures but not the table. The average DSC across all regions was 0.84 and the overall results show that 48 out of 56 regions had a DSC of 0.7 or higher, which is considered excellent agreement. The kappa values for the other regions were between 0.6-0.7, except for Pineal Gland with an average DSC of 0.227. The Spearman correlation between average DSC and volume was 0.703 including Pineal Gland and 0.687 excluding Pineal Gland, which indicates that larger regions are more likely to get a high DSC value compared to smaller regions.

Table 4. Accuracy results for the three accuracy validation subjects. Volume is calculated as the average volume across the three subjects.

Region	Subject 1 DSC	Subject 2 DSC	Subject 3 DSC	Avg DSC	Median DSC	Volume (mm^3)
L Orbitofrontal GM	0.966	0.963	0.954	0.961	0.963	77349
R Orbitofrontal GM	0.938	0.963	0.953	0.951	0.953	78170
L Precentral Gyrus	0.845	0.848	0.747	0.813	0.845	12099
R Precentral Gyrus	0.769	0.849	0.754	0.791	0.769	12987
L Lateral Temporal GM	0.912	0.920	0.923	0.918	0.920	49152
R Lateral Temporal GM	0.930	0.906	0.930	0.922	0.930	47849
L Medial Temporal GM	0.714	0.701	0.772	0.729	0.714	11012
R Medial Temporal GM	0.728	0.715	0.775	0.739	0.728	10733
L Lateral Occipitotemporal Gyrus	0.661	0.645	0.741	0.682	0.661	7149
R Lateral Occipitotemporal Gyrus	0.680	0.670	0.736	0.695	0.680	6932
L Occipital GM	0.827	0.866	0.872	0.855	0.866	47035
R Occipital GM	0.872	0.872	0.885	0.876	0.872	46078
L Parietal GM	0.846	0.903	0.871	0.873	0.871	44009
R Parietal GM	0.876	0.906	0.897	0.893	0.897	41518
L Insula	0.923	0.918	0.907	0.916	0.918	7332
R Insula	0.919	0.904	0.899	0.907	0.904	7668
L Cingulate	0.808	0.868	0.858	0.844	0.858	13409
R Cingulate	0.762	0.873	0.863	0.833	0.863	14204
L Caudate Nucleus	0.933	0.924	0.917	0.925	0.924	3257

Table 4. Accuracy results for the three accuracy validation subjects. Volume is calculated as the average volume across the three subjects. (continued)

Region	Subject 1 DSC	Subject 2 DSC	Subject 3 DSC	Avg DSC	Median DSC	Volume (mm^3)
R Caudate Nucleus	0.947	0.935	0.926	0.936	0.935	3309
L Putamen	0.877	0.877	0.867	0.873	0.877	4822
R Putamen	0.880	0.881	0.852	0.871	0.880	4531
L Globus Pallidus	0.728	0.762	0.534	0.674	0.728	1013
R Globus Pallidus	0.734	0.708	0.625	0.689	0.708	998
L Thalamus	0.910	0.917	0.918	0.915	0.917	7585
R Thalamus	0.927	0.914	0.922	0.921	0.922	7841
L Accumbens	0.748	0.673	0.717	0.712	0.717	524
R Accumbens	0.713	0.701	0.752	0.722	0.713	544
L Amygdala	0.778	0.824	0.784	0.795	0.784	2284
R Amygdala	0.706	0.784	0.816	0.768	0.784	2345
L Hippocampus	0.795	0.763	0.740	0.766	0.763	2800
R Hippocampus	0.794	0.776	0.838	0.802	0.794	2852
Brain Stem	0.968	0.957	0.963	0.963	0.968	28416
L Cerebellum	0.974	0.963	0.971	0.969	0.971	61707
R Cerebellum	0.984	0.971	0.973	0.976	0.973	59534
L Frontal WM	0.976	0.976	0.968	0.973	0.976	68166
R Frontal WM	0.976	0.976	0.972	0.975	0.976	69144
L Temporal WM	0.943	0.955	0.901	0.933	0.943	30731
R Temporal WM	0.931	0.917	0.928	0.926	0.928	30240
L Occpital WM	0.882	0.918	0.843	0.881	0.882	26713
R Occipital WM	0.908	0.876	0.885	0.890	0.885	26577
L Parietal WM	0.924	0.931	0.898	0.917	0.931	37345
R Parietal WM	0.943	0.923	0.916	0.927	0.923	36472
L External Capsule	0.666	0.684	0.697	0.682	0.684	2378
R External Capsule	0.721	0.660	0.642	0.674	0.660	2272
L Int Caps	0.770	0.833	0.746	0.783	0.770	7311
R Int Caps	0.781	0.817	0.751	0.783	0.781	7500
L Fornix	0.668	0.538	0.599	0.602	0.599	751
R Fornix	0.714	0.646	0.707	0.689	0.707	575
Corpus Callosum	0.912	0.906	0.922	0.913	0.912	13367
L Lateral Ventricle CSF	0.970	0.948	0.964	0.961	0.964	20951
R Lateral Ventricle CSF	0.967	0.982	0.980	0.976	0.980	16298
Third Ventricle CSF	0.924	0.927	0.910	0.920	0.924	3010
Fourth Ventricle CSF	0.965	0.991	0.932	0.963	0.965	1301
Cortical CSF	0.997	0.997	0.997	0.997	0.997	381868
Pineal Gland	0.161	0.121	0.398	0.227	0.161	89

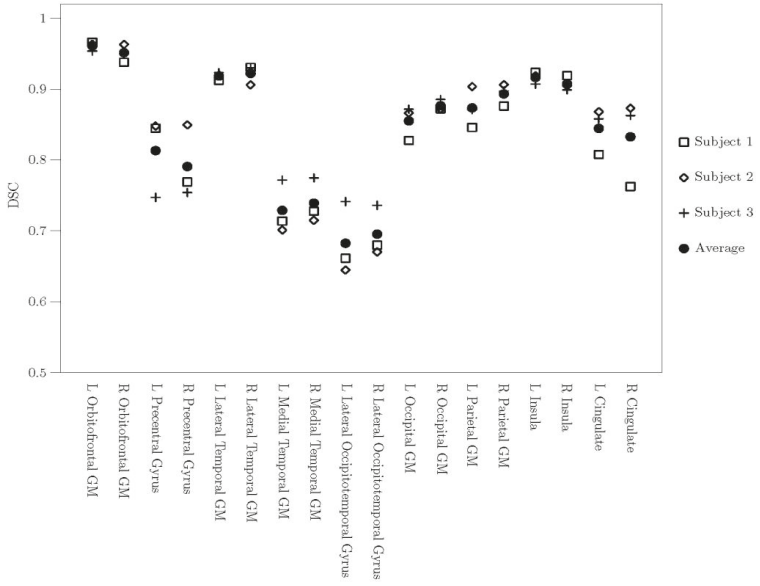

Figure 11. DSC accuracy results for cortical regions.

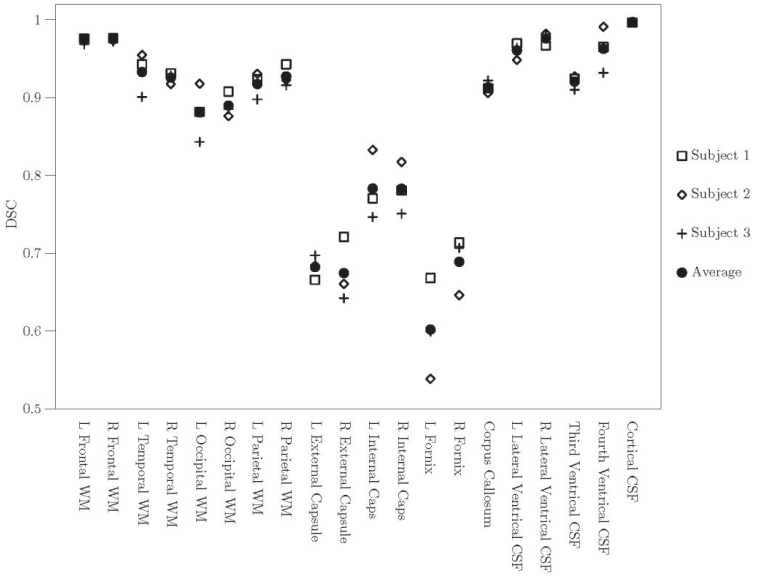

Figure 12. DSC accuracy results for WM and CSF regions.

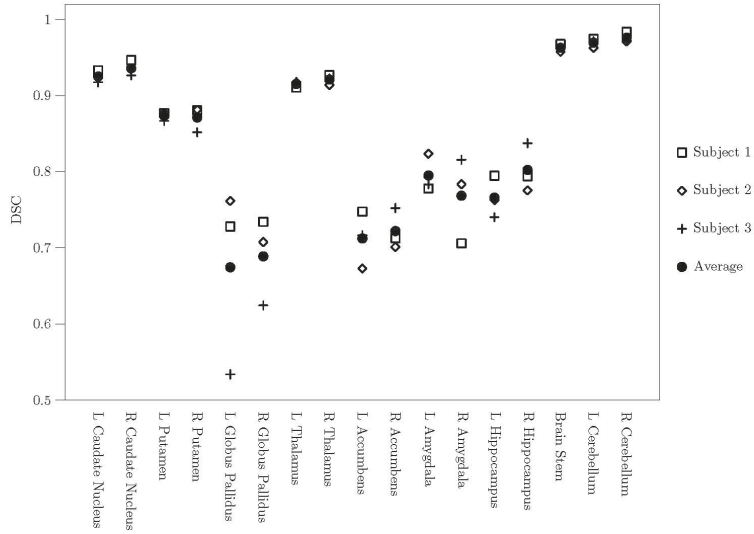

Figure 13. DSC accuracy results for subcortical regions.

DISCUSSION

Validation results

We have developed a template and a regional atlas optimised for old subjects based on automatic segmentation of 314 subjects, where the atlas consists of 56 regions divided into different probabilistic regional tissue atlases. The generation of the template was monitored and showed to converge. A regional segmentation pipeline that uses the template and the probabilistic atlases has also been developed. Furthermore, we validated the different components to assure accuracy and reproducibility. The accuracy testing showed excellent results for 48 out of 56 regions, where accuracy DSC>0.70. Common for all regions with accuracy DSC<0.70 is that they were small in size (less than 6000 mm3) and still had a DSC>0.60 (with an exception of Pineal Gland). The mean accuracy DSC for all regions was 0.84.

For reproducibility, we expect to get a low IOI, defined as IOI < 0:60, and the reproducibility test showed IOI < 0:5 for all regions. This indicates a good level of reproducibility. Also, the mean reproducibility DSC>0.70 for all regions, which demonstrates excellent reproducibility. The results also showed that the mean reproducibility DSC is lower than the median DSC for each region. One reason is that DSC cannot exceed the value 1.0 and is therefore not normally distributed. This means two things; there were more subjects above than below the mean DSC and the subjects below the mean DSC were more spread out. For this reason, the median value is also important to take into consideration.

Whereas accuracy validation showed a strong correlation between volume and mean DSC, the reproducibility test did not show such a relationship. This shows that the level of reproducibility is not related to the size of the regions, which is important for longitudinal studies where the same brain is scanned twice with some time difference.

Comparison with other studies

The AGES atlas covers the whole brain and includes both cortical and subcortical regions. Cortical regions are difficult to objectively define from structural MR images since there is no clear visual border between different lobes. This section is therefore limited to the subcortical regions, which are well defined and should thus be comparable between studies. The DSC measurement is a common method for validation and a comparison of DSC results between studies is given in a review of atlas-based segmentation.[29] Some of the studies use a multi-atlas fusion technique, where each atlas corresponds to a single manually labelled subject.[30-33] In this approach, each atlas is registered to the target image to be segmented and a vote rule decision is applied to get a consensus region for each voxel. Another multi-atlas technique is given by Fischl et al.[34] and Han and Fischl,[35] where an atlas was built based on probabilistic information estimated from a number of manually labelled subjects.

Table 5 compares DSC from these studies with the accuracy results from the AGES pipeline. The regions included are Caudate, Thalamus, Putamen, Globus Pallidus (Pallidum), Hippocampus, and Amygdala. The average DSC in AGES for these regions was 0.83, while the lowest DSC was 0.68 (Globus Pallidus). In the case of Globus Pallidus, there is one accuracy subject that had much lower accuracy DSC than the other two subjects. The other two accuracy subjects had a mean Globus Pallidus DSC of 0.733. This may be due to the reproducibility variance, reflected by the broad range of reproducibility DSC values found for this region, showing that a rescan of the same subject may result in a different result. This is also reflected by the high CV_W value for Globus Pallidus, which was higher than many other regions. Globus Pallidus is a common place for infarcts in old subjects,[36] which may affect both the accuracy and reproducibility results. However, the IOI value for Globus Pallidus was 0.29, which is still considered good.

Common for all studies in this comparison is that their cohorts average age are younger than the AGES cohort. Given the heterogenous nature of an old cohort, it is more challenging to get a good result. The DSC values depend to a high degree on the cohort, the testing data and the definition of the different regions. A quantitative comparison between studies based on DSC is therefore difficult but can at least give an indication to whether or not a particular study gives decent results. One can also notice that not a

single study outperforms the others for all the regions. This comparison demonstrates that the AGES result is on par with other studies of younger populations.

Table 5. Comparison between studies using the DSC metric. The average DSC across hemispheres is first calculated for each subject before calculating each region's median DSC in the AGES study. The regions are: Caudate (CAU), Thalamus (THA), Putamen (PUT), Globus Pallidus (PAL), Hippocampus (HIP), and Amygdala (AMY).

Study	CAU	THA	PUT	PAL	HIP	AMY	Average
AGES study: Avg DSC	0.93	0.92	0.87	0.68	0.78	0.78	0.83
AGES study: Median DSC	0.93	0.92	0.88	0.73	0.87	0.85	0.86
AGES study: CV_w (%)	2.06	1.05	2.04	3.79	1.42	2.78	2.19
AGES study: IOI	0.20	0.13	0.19	0.29	0.12	0.20	0.19
Fiscl et al. (2002)[34]	0.88	0.79	0.71	0.71	0.71	0.79	0.76
Heckerman et al. (2006)[30]	0.90	0.90	0.90	0.80	0.81	0.80	0.85
Han and Fischl (2007)[35]	0.84	0.88	0.85	0.76	0.83	0.75	0.82
Aljabar et al. (2009)[31]	0.88	0.91	0.90	0.82	0.83	0.78	0.85
Artaechevarria et al. (2009)[32]	0.83	0.8	0.87	0.81	0.75	0.72	0.81
Lötjönen et al. (2010)[33]	0.87	0.90	0.90	0.84	0.82	0.77	0.85
Average	0.88	0.88	0.86	0.77	0.79	0.77	0.83

Study strengths and limitations

The main results in this study show that we have a robust probability atlas and regional segmentation pipeline that have been validated thoroughly for an old cohort. This has been done through a comprehensive validation by testing both the accuracy and reproducibility of the segmentation procedure. The template and the probability atlases are based on 314 subjects and 56 regions. All this taken together make this atlas and template unique.

The design goal for the AGES atlas was to create a non-linear symmetrical template and corresponding atlas that could be used for both regional segmentation and as a reference atlas in MNI space. This required a single atlas in linear MNI space, rather than a multiatlas fusion technique used by other studies. However, by separating the atlas into different tissue-atlases, the inter-subject variability is well contained. Another design decision was whether or not to manually label the AGES template. Since the AGES template is an average brain, it would be difficult to manually label regions that are fuzzy. It was therefore decided to manually label individual subjects that are then warped onto the template to create an initial atlas. The template includes a few MCI and dementia cases to make the template more representative for an elderly population in general. Note that although the template was constructed in linear MNI space, its non-linear features are based on the AGES-Reykjavik cohort. Thus, no spatial distortions were introduced by

the linear registration to MNI space. The template and atlas were made symmetrical to prevent bias due to asymmetrical variability across subjects. There are some limitations in this study that need to be highlighted. Only seven subjects were manually labelled, of which three were used for accuracy testing. Many more subjects would be necessary to make the accuracy tests statistically meaningful. Unfortunately, manual labelling of the whole brain is a time consuming process and the study was for this reason limited to seven subjects. Given this limitation, a choice had to be made between the number of subjects to be used for constructing the initial atlas and for accuracy testing. Here, one could choose a jackknife procedure and use six subjects for creating the atlas and keep one for testing, giving seven different atlases with one test subject for each atlas. However, that would only give one accuracy test subject per atlas, making it difficult to judge the performance of each atlas. Instead, four subjects were used for creating the initial atlas and three subjects were used for accuracy testing, which was thought to be a fair balance. The accuracy results have to be considered with this in mind; still, it gives some insight into how well a few cases are regionally segmented by using this atlas.

The ageing brains are affected by large morphological changes. In that aspect, an old cohort is more heterogeneous than a young cohort. Since the calculations of the accuracy DSC metric are limited to only three subjects, more subjects would absolutely be needed to represent every aspect of the ageing brain. To compensate for this, repeated scans from 31 subjects were also included to measure reproducibility using DSC and CV. It is important to note that the DSC between the accuracy and reproducibility are not comparable. Both the reproducibility of the automatic tissue segmentation and the regional segmentation are tested using the repeated scans. However, in the manual labelling procedure, the automatic tissue segmentations were used as guidelines and the labeller could choose to follow these guidelines or draw freehand. The guidelines were often followed for the border between GM and WM in cortical regions, but less so in subcortical regions. By following these guidelines, it was thought that this would increase the quality of the atlas in regions where the automatic tissue segmentation did a good job (as judged by the labeller). This however inevitably increased the DSC slightly in regions where the border between GM and WM was defined by the automatic tissue segmentation. This is thus not a limitation of the construction of the atlas, but it is a limitation in the accuracy testing.

A few regions had accuracy DSC<0.70. This indicates less accurate results for these regions. However, most of these regions were close to 0.70 and their results should still be considered good. Only Pineal Gland had a low accuracy, which is due to its small size making it a difficult region to segment automatically. For reproducibility testing, the range of the DSC values per region indicates that a few subjects are below DSC<0.70

for some regions. Nevertheless, besides Pineal Gland, all regions had a DSC> 0:70 for the first quartile demonstrating a good reproducibility for most subjects. This study did not include a comparison with another atlas using the same subjects. Given that the test subjects are labelled specifically for this atlas, it would be difficult to use the accuracy test subjects to test a different atlas with different regions.

CONCLUSIONS

There are many freely available atlases that can be used for brain studies, but most are generated from relatively young individuals. The results of this study show evidence of a robust atlas and regional segmentation pipeline for an ageing population. This publicly available regional atlas and corresponding template is the first old cohort brain atlas that has been extensively validated and which covers the whole brain.

REFERENCES

1. Seitz RJ, Bohm C, Greitz T, Roland PE, Eriksson L, Blomqvist G, Rosenqvist G, Nordell B. Accuracy and precision of the computerized brain atlas programme for localization and quantication in positron emission tomography. *J Cereb Blood Flow Metab.* 1990; 10(4), 443-57.
2. Roland PE, Graufelds CJ, Hlin JW, Ingelman L, Andersson M, Ledberg A, Pedersen J, Akerman S, Dabringhaus A, Zilles K. Human brain atlas: For high-resolution functional and anatomical mapping. *Human brain mapping.* 1994; 1(3): 173-84.
3. Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J. A probabilistic atlas of the human brain: theory and rationale for its development. the international consortium for brain mapping (icbm). *Neuroimage*. 1995; 2(2):89-101.
4. Toga AW, Thompson PM. The role of image registration in brain mapping. *Image Vis Comput*. 2001;19 (1-2), 3-24.
5. Thompson PM, Mega MS, Vidal C, Rapoport JL, Toga AW. Detecting disease-specific cpatterns of brain structure using cortical pattern matching and a population-based probabilistic brain atlas. *Inf Process Med Imaging.* 2001; 2082: 488-501.
6. Fonov V, Evans A, McKinstry R, Almli C, Collins D. Unbiased nonlinear average ageappropriate brain templates from birth to adulthood. *NeuroImage*. 2009; Supplement 1 (0), S102 EP.
7. Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL, Group BDC. Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage*. 2011;54 (1):313-27.
8. Ashburner J, Friston K. Multimodal image coregistration and partitioning- a unified framework. *Neuroimage*.1997;6(3):209-17.
9. Collins D, Zijdenbos A, Baare W, Evans A. Animal+insect: Improved cortical structure segmentation. 1999
10. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*. 2002;15(1):273-89.
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 et al. A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (icbm). *Philos Trans R Soc Lond, B, Biol Sci*. 2001;356(1412):1293-322.
12. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, J Johansen-Berg, H, Bannister PR, Luca MD, DrobnjakI, Flitney DE et al. Advances in functional and structural mr image analysis and implementation as fsl. NeuroImage. 2004;23 Suppl 1: 208-19.
13. Shattuck DW, Mirza M, Adisetiyo V, Hojatkashani C, Salamon G, Narr KL, Poldrack RA, Bilder RM, Toga AW. Construction of a 3d probabilistic atlas of human cortical structures. Neuroimage. 2008;39(3):1064-80.
14. Machilsen B, d'Agostino E, Maes F, Vandermeulen D, Hahn HK, Lagae L, Stiers P. Linear normalization of mr brain images in pediatric patients with periventricular leukomalacia. *Neuroimage*.2007;35(2):686-97.
15. Mega MS, Dinov ID, Mazziotta JC, Manese M, Thompson PM, Lindshield C, Moussai J, Tran N, Olsen K, Zoumalan CI et al. Automated brain tissue assessment in the elderly and demented population: construction and validation of a sub-volume probabilistic brain atlas. *Neuroimage*. 2005;26(4):1009-18.
16. Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, Frackowiak RS. A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*. 2001;14(1 Pt 1):21-36.

17. Levy-Cooperman N, Ramirez J, Lobaugh NJ, Black SE. Misclassified tissue volumes in alzheimer disease patients with white matter hyperintensities: importance of lesion segmentation procedures for volumetric analysis. *Stroke*. 2008;39(4):1134-41.
18. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, Thorgeirsson G, Aspelund T, Garcia ME, Cotch MF et al. Age, gene/environment susceptibility-reykjavik study: multidisciplinary applied phenomics. *Am J Epidemiol*. 2007;165(9):1076-87.
19. Sigurdsson S, Aspelund T, Forsberg L, Fredriksson J, Kjartansson O, Oskarsdottir B, Jonsson PV Eiriksdottir G, Harris TB, Zijdenbos A, van Buchem MA, Launer LJ. Gudnason V. Brain tissue volumes in the general population of the elderly: the ages-reykjavik study. *Neuroimage*. 2012;59(4): 3862-70.
20. Sled J, Zijdenbos A, Evans A. A comparison of retrospective intensity non-uniformity correction methods for mri. Information processing in medical imaging. 1997.
21. Collins DL, Neelin P, Peters TM., Evans AC. Automatic 3d intersubject registration of mr volumetric data in standardized talairach space. *Journal of Computer Assisted Tomography.* 1994;18(2):192-205.
22. Zijdenbos, A. P., Forghani, R., Evans, A. C. Automatic "pipeline" analysis of 3-d mri data for clinical trials: application to multiple sclerosis. *IEEE transactions on medical imaging*. 2002;21(10):1280-91.
23. Smith SM. Fast robust automated brain extraction. *Human brain mapping.* 2002;17(3): 143-55.
24. Collins D, Evans A. Animal: Validation and applications of nonlinear registration-based segmentation. *Int J Pattern Recogn*. 1997;11(8):1271-1294.
25. Huo J, Alger J, Kim H, Brown M Okada K Pope W Goldin, J. Between-scanner and between-visit variation in normal white matter apparent diffusion coeffcient values in the setting of a multicenter clinical trial. *Clin Neuroradiol*. 2015.
26. Grech-Sollars M, Hales PW, Miyazaki K, Raschke F, Rodriguez D, Wilson M. Gill SK, Banks T, Saunders DE, Clayden JD, et al. Multi-centre reproducibility of diffusion mri parameters for clinical sequences in the brain. *NMR Biomed.* 2015;28(4):468-85.
27. Quan H, Shih W. Assessing reproducibility by the within-subject coeffcient of variation with random effects models. *Biometrics.* 1996;52(4):1195-1203.
28. Harris EK. Effects of intra- and interindividual variation on the appropriate use of normal ranges. *Clin Chem*. 1974;20(12):1535-42.
29. Cabezas M, Oliver A, Llado X, Freixenet J, Cuadra MB. A review of atlas-based segmentation for magnetic resonance brain images. *Comput Methods Programs Biomed*. 2011;104(3):58-77.
30. Heckemann RA, Hajnal JV, Aljabar P, Rueckert D, Hammers A. Automatic anatomical brain mri segmentation combining label propagation and decision fusion. *Neuroimage*. 2006;33(1):115-26.
31. Aljabar P, Heckemann RA, Hammers A, Hajnal JV, Rueckert D. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage*. 2009;46(3):726-38.
32. Artaechevarria, X, Munoz-Barrutia A, de Solorzano CO. Combination strategies in multiatlas image segmentation: application to brain mr data. *IEEE Trans Med Imaging*. 2009;28(8):1266-77.
33. Löotjöonen JM, Wolz R, Koikkalainen JR, Thurfjell L, Waldemar G, Soininen H, Rueckert D. Initiative ADN: Fast and robust multi-atlas segmentation of brain magnetic resonance images. *Neuroimage*. 2010;49(3): 2352-65.
34. Fischl B, Salat DH, Busa E, Albert M, Dieterich, M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron.* 2002;33(3):341-55.

35. Han X, Fischl B. Atlas renormalization for improved brain mr image segmentation across scanner platforms. *IEEE Trans Med Imaging.* 2007;26(4):479-86.
36. Mori E. Impact of subcortical ischemic lesions on behavior and cognition. *Ann N Y Acad Sci*. 2002;977:141-8.