NeuroImage 52 (2010) 69–85

Contents lists available at ScienceDirect

NeuroImage

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

Reproducibility of thalamic segmentation based on probabilistic tractography

Catherine Traynor [a], Rolf A. Heckemann [c,d], Alexander Hammers [c,d], Jonathan O'Muircheartaigh [a], William R. Crum [b], Gareth J. Barker [b], Mark P. Richardson [a,*]

[a] King's College London Institute of Psychiatry, Department of Clinical Neuroscience, England, UK
[b] King's College London Institute of Psychiatry, Department of Neuroimaging, England, UK
[c] MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, England, UK
[d] The Neurodis Foundation (Fondation Neurodis), Lyon, France

ARTICLE INFO

Article history:
Received 19 October 2009
Revised 6 April 2010
Accepted 7 April 2010
Available online 14 April 2010

ABSTRACT

Reliable identification of thalamic nuclei is required to improve targeting of electrodes used in Deep Brain Stimulation (DBS), and for exploring the role of thalamus in health and disease. A previously described method using probabilistic tractography to segment the thalamus based on connections to cortical target regions was implemented. Both within- and between-subject reproducibility were quantitatively assessed by the overlap of the resulting segmentations; the effect of two different numbers of target regions (6 and 31) on reproducibility of the segmentation results was also investigated. Very high reproducibility was observed when a single dataset was processed multiple times using different starting conditions. Thalamic segmentation was also very reproducible when multiple datasets from the same subject were processed using six cortical target regions. Within-subject reproducibility was reduced when the number of target regions was increased, particularly in medial and posterior regions of the thalamus. A large degree of overlap in segmentation results from different subjects was obtained, particularly in thalamic regions classified as connecting to frontal, parietal, temporal and pre-central cortical target regions.

© 2010 Elsevier Inc. All rights reserved.

Introduction

Deep brain stimulation (DBS) of thalamic nuclei is being developed as a treatment for drug-resistant epilepsy (Theodore and Fisher, 2004) and is widely used in the treatment of movement disorders (Koller et al., 1997). Reliable identification of thalamic nuclei has the potential to improve treatment success, by improving the accuracy with which DBS electrodes are targeted. Currently, targeting of DBS electrodes uses qualitative magnetic resonance images in combination with thalamic atlases, for example, the Morel thalamic atlas (Morel et al., 1997) or Schaltenbrand and Wahren atlas (Schaltenbrand and Wahren, 1977). Typically, the atlases used are based on a small number of subjects and do not take into account inter-subject variability. Furthermore, atlases may present data from different subjects sectioned in three orthogonal planes, and there is good evidence that these orthogonal sections do not correspond sufficiently to provide reliable targeting (Nowinski et al., 2008). New methods which reliably and accurately identify individual nuclei within the thalamus are required to improve targeting in DBS, and will also have great potential for examination of the integrity of thalamic nuclei across a range of diseases.

Recent work has shown that thalamic sub-regions, thought to correspond to groups of thalamic nuclei, can be identified from diffusion imaging data at the single-subject level using connectivity-based segmentation (Behrens et al., 2003a; Johansen-Berg et al., 2005). Each thalamus contains different nuclei that can be distinguished histologically by their characteristic cyto- and myelo-architecture (Morel et al., 1997). Currently, however, these nuclei cannot be identified on conventional, qualitative MR images due to insufficient contrast in deep grey matter structures. The different nuclei are known to have extensive connections to specific regions of the ipsilateral cerebral cortex (as well as reciprocal connections to each other). Using a probabilistic tractography method, Behrens et al. (2003a) segmented the thalamus based on connections between thalamic voxels and ipsilateral cortical target regions but carried out only a qualitative assessment of the between-subject reproducibility of the method. Segmentation of the thalamus was carried out for eight subjects using four large cortical target regions. For six of the eight subjects, the resulting segmentations were found to be very similar. The diffusion data for the remaining two subjects were of lower quality and hence produced a poorer segmentation. No quantitative assessment was carried out of the reproducibility of the method for a single subject scanned multiple times. The results have been promising, producing thalamic segmentations which appear to be reproducible across subjects (Behrens et al., 2003a; Johansen-Berg et al., 2005), but detailed quantitative assessment of within- and between-subjects reliability has not yet been attempted.

* Corresponding author. P043 Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK. Fax: +44 20 7848 5440.
E-mail address: mark.richardson@kcl.ac.uk (M.P. Richardson).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.04.024

Table 1
Cortical target regions.

Region label	Cortical region	Region label	Cortical region
1	Hippocampus	17	Straight gyrus (gyrus rectus)
2	Amygdala	18	Anterior orbital gyrus
3	Anterior temporal lobe, medial part	19	Inferior frontal gyrus
4	Anterior temporal lobe, lateral part	20	Superior frontal gyrus
5	Gyri parahippocampalis et ambiens	21	Postcentral gyrus
6	Superior temporal gyrus, posterior part	22	Superior parietal gyrus
7	Medial and inferior temporal gyri	23	Lingual gyrus
8	Lateral occipitotemporal gyrus (gyrus fusiformis)	24	Cuneus
9	Insula	25	Medial orbital gyrus
10	Occipital lobe	26	Lateral orbital gyrus
11	Cingulate gyrus, anterior part	27	Posterior orbital gyrus
12	Cingulate gyrus, posterior part	28	Subgenual frontal cortex
13	Middle frontal gyrus	29	Subcallosal area
14	Posterior temporal lobe	30	Pre-subgenual frontal cortex
15	Parietal lobe	31	Superior temporal gyrus, anterior part
16	Precentral gyrus		

In the study reported here, a quantitative assessment of the reproducibility of the connectivity-based method of thalamic segmentation was carried out, both within and between subjects. Initially, we examined reproducibility of the thalamic segmentation method in a single subject. We approached this problem using multiple re-analyses of the same single dataset, varying the starting conditions of the probabilistic tractography algorithm. This was followed by an examination of multiple datasets that had been independently acquired from the same subject. Subsequently we examined the reproducibility of thalamic segmentation between subjects in a group of normal subjects. Additionally, we investigated whether reproducibility of our findings was a function of the number of cortical targets, and whether the degree of contrast in the structural imaging data influenced the definition of cortical and thalamic regions.

Methods

Data acquisition

Sixteen DTI datasets were acquired for a single healthy subject (male, 24 years). These data were acquired in 12 sessions over a period of eight months. In four of the sessions, two DTI datasets were obtained (datasets 1 and 2, 3 and 4, 5 and 6, 15 and 16). The remaining eight datasets were acquired in separate sessions. A 3T, GE Signa HDx system (General Electric, Waukesha, WI, USA), with actively shielded magnetic field gradients and a maximum gradient amplitude of 40 mT m^{-1}, was used. The acquisition was based on the protocol described by Jones et al. (2002b). The body coil was used for RF transmission, and an 8-channel head coil for signal reception, allowing a parallel imaging (ASSET) speed-up factor of two. Each volume was acquired using a multi-slice peripherally-gated doubly-refocused spin-echo EPI sequence. Sixty contiguous, near-axial slice locations with 2.4 mm isotropic voxels were acquired. The echo time was 104.5 ms. The effective repetition time varied between sessions from 15 to 20 RR intervals. The maximum diffusion weighting was 1300 s m^{-2}. At each slice location the acquisition comprised four images with no diffusion gradients applied and 32 diffusion-weighted images in which the gradient directions were uniformly distributed in space. The scanning time for DTI data acquisition at each session was approximately 10 min.

In addition to the DTI datasets, a structural image was required on which the cortical target regions and thalamic seed regions could be identified. Two structural images were acquired; a coronal inversion-recovery prepared fast spoiled gradient echo (IR-FSPGR) image and a sagittal T1 map. The IR-FSPGR image was acquired using the following parameters: TI = 450 ms, TR = 6.976 ms, TE = 2.812 ms, excitation flip angle 20°, 256 × 256 × 196 matrix. The T1 map was obtained using the DESPOT-HIFI method described by Deoni (2007a). To generate this T1 map, two SPGR images (TR = 7.932 ms, TE = 3.752 ms, excitation flip angles 18° and 4°, 256 × 256 × 160 matrix) and one IR-FSPGR image (TI = 450 ms, TR = 7.684 ms, TE = 3.752 ms and flip angle 5°, acquired with voxel dimensions double those of the SPGR images in antero-posterior and left–right directions, then zero-padded to the same size before Fourier transformation) were acquired and processed using the method described by Deoni (2007a) and ImageJ plugins (downloaded from http://users.fmrib.ox.ac.uk/~sdeoni/Analysis%20Code.html). The IR-FSPGR image acquired as part of the DESPOT method was not used as the second structural image because it was acquired at a lower resolution and was not optimised for contrast.

The coronal IR-FSPGR image and the T1 map both had an in-plane field of view of 28 cm, and a slice thickness (voxel size in the second, through plane, phase encode direction) of 1.1 mm, leading to a final voxel size of 1.1 mm × 1.1 mm × 1.1 mm. Each of these structural images was used separately as the basis for generating the seed and target regions required for the thalamic segmentation.

For the 18 additional healthy volunteers, DTI data and structural images were again acquired using the parameters specified above, in a single session per subject. The subjects included 11 female subjects (age range 25–47 years) and 7 male subjects (age range 23–43 years).

The study was approved by the local research ethics committee, and informed consent obtained from all participants.

Pre-processing

Following acquisition of the DTI data, initial processing aimed to reduce the impact of distortions and movement artefacts in the images. This was achieved using in-house software, which combined steps for registering the DTI data together using FLIRT (Jenkinson and Smith, 2001) and masking the resultant images using the method described by Jones et al. (2002a). This involved performing brain extraction on one of the non-diffusion weighted images using Brain Extraction Tool (BET) (Smith, 2002). A binary mask was then created from the processed non-diffusion weighted image and used to mask all other images in the diffusion dataset. In addition, two files containing information about the gradient directions (bvecs) and their corresponding b-value (bvals) were created using in-house software.

Thalamic seed masks and cortical target regions

For segmentation of the thalamus, connections from the thalamus to different cortical target regions were investigated. This required the creation of thalamic seed masks and target masks for all cortical

Table 2
Regions from Table 1 included in the six larger cortical target regions.

Large cortical regions	Regions from Table 1 included in larger regions
Frontal	11, 13, 17, 18, 19, 20, 25, 26, 27, 28, 29, 30
Occipital	10, 23, 24
Parietal	12, 15, 22
Post-Central	21
Pre-Central	16
Temporal	1, 2, 3, 4, 5, 6, 7, 8, 14, 31

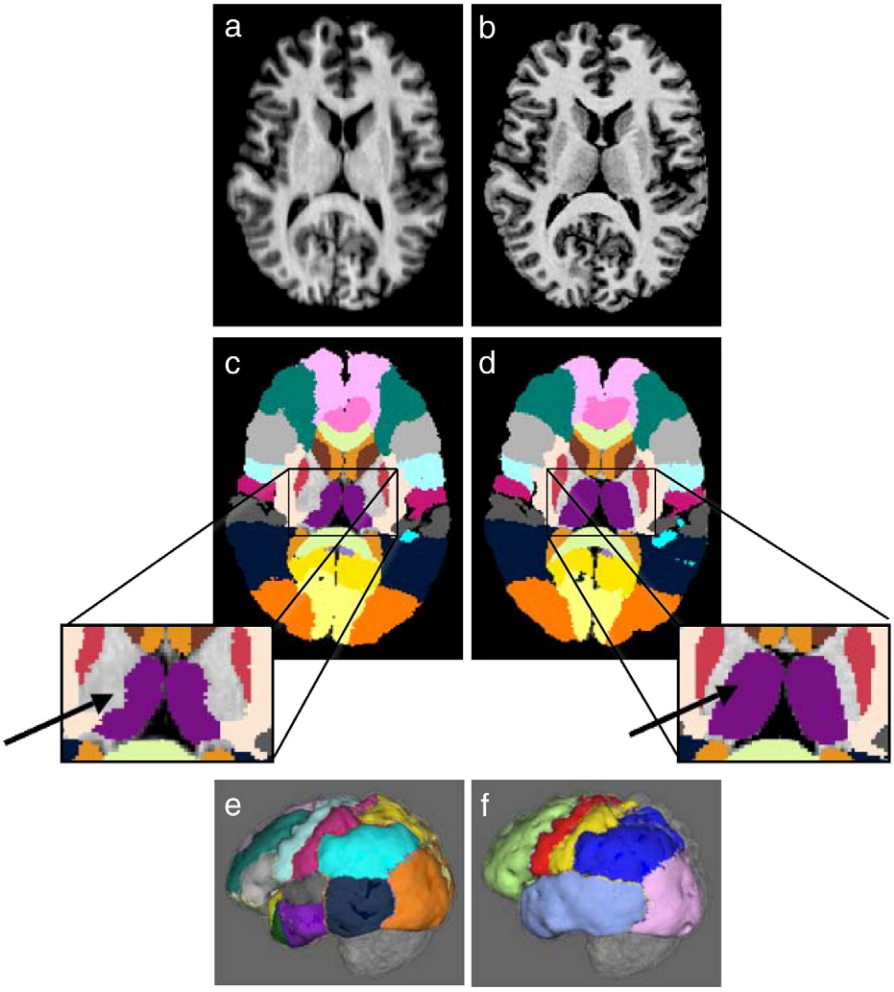

Fig. 1. An axial slice from (a) the coronal IR-FSPGR structural image and (b) the inverted contrast T1 map, and the corresponding cortical and sub-cortical regions generated based on these two images ((c) and (d), respectively). The cortical regions generated using the two structural images are very similar. However, the thalamic regions are considerably different, as shown in the enlarged sections, with the lateral edge of the thalamus excluded from the region identified from the IR-FSPGR image (c). (e) and (f) show 31 and six cortical target regions on the surface of the brain. (Note: for images (a) and (b) the IR-FSPGR image was registered to the inverted contrast T1 map. The high contrast between thalamus and adjacent white matter is still apparent if the inverted contrast T1 map is registered to the IR-FSPGR, see Supplementary Figure 1).

regions of interest. We used multi-atlas label propagation and decision fusion as described by Heckemann et al. (2006) to create all these masks. A previously collected set of 30 atlases (IR-FSPGR images from normal controls, which had been manually segmented into 83 cortical and sub-cortical regions; demographic details in Gousias et al. (2008) and additional protocols in Hammers et al. (2003)) was used to segment the image data collected in the current study. MR images of each atlas were paired with each subject for non-rigid registration. The registration procedure was based on hierarchical free-form deformation via evenly spaced control points (20, 10 and 5 mm spacing), blended using B-splines (Rueckert et al., 1999). The resulting spatial transformations were used to propagate the atlas label sets into the space of the target subject. This resulted in 30 segmentations for each subject, which were consolidated using vote-rule decision fusion, where each voxel is assigned a unique label which corresponds to the majority of label assignments to this voxel from the individual segmentations. The resulting label set contained all 83 regions. In this study, connectivity-based thalamic segmentation was carried out using two different cortical target region sets: in the first analysis six target regions per hemisphere were used; this was then followed by a second analysis using 31 target regions per hemisphere (see Table 1). The 31 targets included all cortical regions identified using the segmentation method described above. For each hemisphere, these regions were extracted from the original image, producing separate images for each target region. Thirty regions (excluding the insula) were combined, as shown in Table 2, to produce the six larger target regions. Only six large regions were used instead of the seven used by Behrens et al. (2003a), as our cortical segmentation did not include the subdivision of the pre-central area into M1 and pre-motor cortical regions.

The cortical regions generated contained both grey matter (GM) and white matter (WM). To obtain cortical masks of GM only, the chosen structural image was segmented into partial volume maps of GM, WM and cerebrospinal fluid (CSF) using FAST (Zhang et al., 2001). A threshold of 35% was then applied to the GM partial-volume image, and the resulting image was binarised. Each of the cortical regions of interest was then used to mask this binary image, creating 31 regions of GM for each hemisphere in the form required for probabilistic tractography.

For the first subject, the IR-FSPGR image and T1 map were used separately as the basis for obtaining cortical and thalamic regions. Before being segmented, the contrast of the T1 map was inverted. This was required to produce an image with contrast similar to a T1-weighted image, as the method for finding cortical and sub-cortical regions had previously been tested using this type of image (Heckemann et al., 2006). Overlap of the cortical regions identified from each of the two structural images was calculated using the equation for Total Accumulated Overlap (TAO; see below). Additionally, overlap of the thalamus

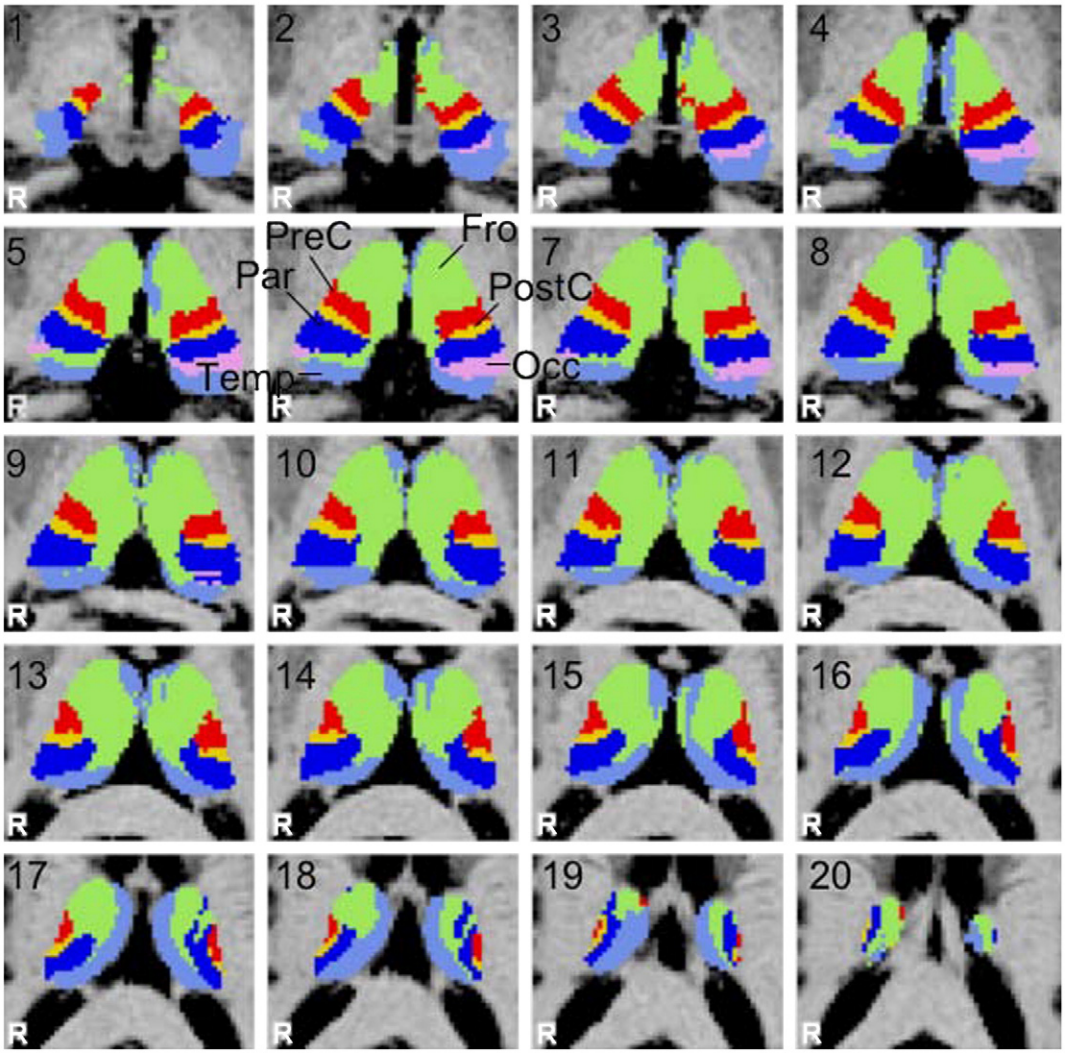

Fig. 2. 20 axial slices through the thalami, from inferior (Slice 1) to superior (Slice 20) showing the regions identified using hard segmentation carried out once in the first subject. The results of the hard segmentation are overlaid on the inverted-contrast T1 map of the same subject. Six target regions were used, including frontal (Fro, green), occipital (Occ, pink), parietal (Par, blue), post-central (PostC, yellow), pre-central (PreC, red) and temporal (Temp, blue) regions. Segmentations of the right and left thalami were carried out independently.

regions identified from the two structural images was assessed using the Overlap By Pairs (OBP) value described below. Regions identified from the two images were visually compared. Thalamic regions generated were also compared to the probabilistic maps of the thalamus shown in Ahsan et al. (2007). Based on this visual inspection, one structural image and its associated regions were chosen for use in the probabilistic tractography algorithm.

The cortical and thalamic regions created using the method described above were all defined in subject-specific, native space using the untransformed structural image (structural space). For probabilistic tractography, further processing was carried out in the lower-resolution diffusion space defined by the DTI scan parameters. A transformation matrix between the structural and diffusion spaces was required for use in the probabilistic tractography process. This was obtained using FLIRT (Jenkinson and Smith, 2001) with the chosen structural image (IR-FSPGR or inverted-contrast T1 map) used as the reference image and a volume from the DTI acquisition with no diffusion weighting used as the input image. Both of these images

Table 3
OBL values and average volumes of regions classified as connecting to each of the six cortical target regions (single subject, 16 repetitions of probabilistic tractography analysis carried out on the same dataset). As the volume of each region was very consistent across the segmentation results, OBL value calculated using $\alpha_i = 1/V_i$ was the same as using $\alpha_i = 1$.

Region	Left thalamus				Region	Right thalamus			
	Overlap (OBL)		Volume / mm^3			Overlap (OBL)		Volume (mm^3)	
	$\alpha_i = 1$	$\alpha_i = 1/V_i$	Mean	sd		$\alpha_i = 1$	$\alpha_i = 1/V_i$	Mean	sd
Frontal	0.959	0.959	4180.8	10.4	Frontal	0.972	0.972	4544.0	9.2
Occipital	0.854	0.854	368.4	7.4	Occipital	0.643	0.643	57.4	5.4
Parietal	0.953	0.953	1700.7	8.1	Parietal	0.977	0.977	1770.1	5.0
Post-Central	0.957	0.957	377.2	3.4	Post-Central	0.976	0.976	503.3	3.0
Pre-Central	0.937	0.937	920.0	6.7	Pre-Central	0.987	0.987	818.4	2.0
Temporal	0.947	0.947	2178.7	8.4	Temporal	0.944	0.944	2031.7	8.8

Table 4
Fuzzy overlap OBL values for regions classified as connecting to each of the six cortical target regions (single subject, 16 repetitions of probabilistic tractography analysis carried out on the same dataset).

Region	Left thalamus		Region	Right thalamus	
	Fuzzy overlap (OBL)			Fuzzy overlap (OBL)	
	$\alpha_i = 1$	$\alpha_i = 1/V_i$		$\alpha_i = 1$	$\alpha_i = 1/V_i$
Frontal	0.956	0.956	Frontal	0.969	0.969
Occipital	0.854	0.854	Occipital	0.727	0.727
Parietal	0.940	0.940	Parietal	0.963	0.963
Post-Central	0.924	0.924	Post-central	0.961	0.961
Pre-Central	0.936	0.936	Pre-central	0.976	0.976
Temporal	0.937	0.937	Temporal	0.943	0.943

were brain-extracted before being used in FLIRT (Smith, 2002). The default options of 12 degrees of freedom, the mutual information cost function and normal search with tri-linear interpolation were used, as suggested on the FSL website (http://www.fmrib.ox.ac.uk/fsl/fdt/fdt_reg.html).

Probabilistic tractography

Probabilistic tractography was carried out using FMRIB's Diffusion Toolbox (FDT v2.0, http://www.fmrib.ox.ac.uk/fsl/fdt/index.html) in FSL v4.1.1. The "probtrackx" code was modified (Saad Jbabdi, FMRIB, personal communication) to ensure that paths were stopped as they entered the termination mask, instead of when leaving the termination mask; the modified code has since been made the standard in publicly released versions of FSL from 4.1.3. Initially, "bedpostx" was used to determine the diffusion parameters at each voxel. Only one fibre was modelled per voxel, as previous work had shown that using a multi-fibre approach (to model crossing fibres) did not affect the segmentation of the thalamus achieved using the probabilistic tractography algorithm (Behrens et al., 2007). Additionally, in the work by Behrens et al. (2007) 60 diffusion encoding directions were used; in the work presented here, only 32 directions were used. A burn-in of 5000 was chosen, instead of the default value of 1000, to ensure convergence of the Markov chains from which the posterior distributions of the parameters of interest were sampled. The default value for "weight" was used.

Following bedpostx, "probtrackx" was run using a single seed mask of the thalamus and cortical classification targets. In this processing step, 5000 streamlines were followed from each voxel in the thalamus mask. For probtrackx, a cortical termination mask was created by combining the GM-only cortical target regions. This termination mask was required to ensure that paths stopped as soon as they entered the cortex, and to prevent paths from passing from one target region into another. The images of the mask and target regions required by probtrackx were input in structural space. Processing was carried out independently for each cerebral hemisphere, with connections from each thalamus to the ipsilateral cortical target regions being

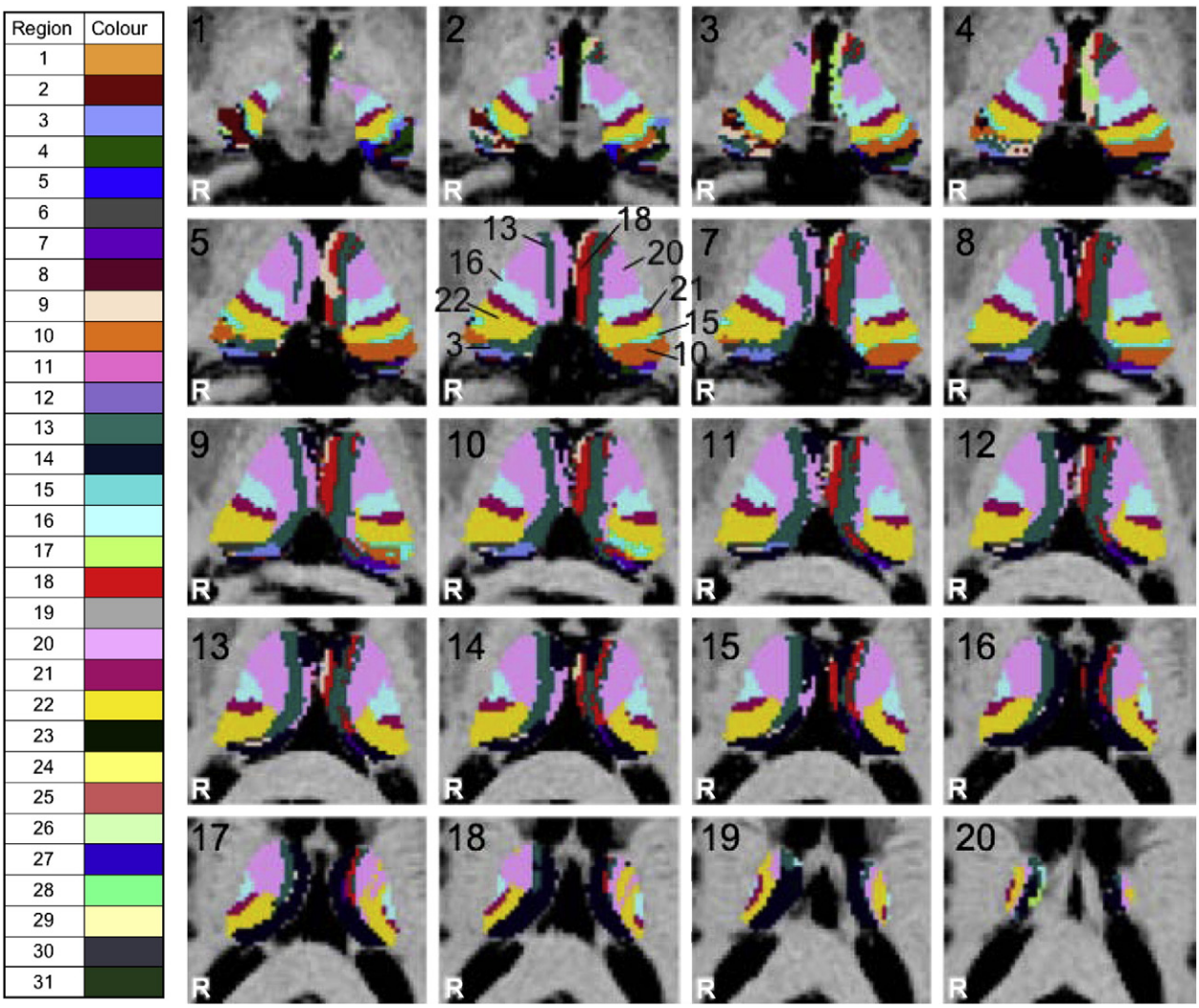

Fig. 3. 20 axial slices, from inferior (Slice 1) to superior (Slice 20,) showing the results of hard segmentation of the thalamus of one dataset of the first subject, overlaid on the inverted-contrast T1 map. Segmentations of the right and left thalami were carried out independently using 31 ipsilateral cortical target regions.

Table 5
OBL values and average volumes of regions classified as connecting to each of the 31 cortical target regions (single subject, 16 repetitions of probabilistic tractography analysis).

Region	Left thalamus				Region	Right thalamus			
	Overlap (OBL)		Volume (mm^3)			Overlap (OBL)		Volume (mm^3)	
	$\alpha_i=1$	$\alpha_i=1/V_i$	Mean	sd		$\alpha_i=1$	$\alpha_i=1/V_i$	Mean	sd
1	0.098	0.089	3.9	1.7	1	0.733	0.777	3.2	1.0
2	0.713	0.713	119.6	3.6	2	0.744	0.744	255.9	8.5
3	0.771	0.772	103.1	5.7	3	0.840	0.840	198.5	3.5
4	0.777	0.777	129.6	6.4	4	0.000	0.000	0.1	0.3
5	0.726	0.726	58.2	2.3	5	0.000	0.000	0.0	0.0
6	0.000	0.000	0.2	0.5	6	0.845	0.845	96.7	3.0
7	0.869	0.869	160.3	3.9	7	0.000	0.000	0.0	0.0
8	0.000	0.000	0.0	0.0	8	0.000	0.000	0.2	0.5
9	0.744	0.744	255.5	8.8	9	0.648	0.648	167.0	7.6
10	0.860	0.860	470.3	5.6	10	0.688	0.688	85.5	3.5
11	0.000	0.000	0.0	0.0	11	0.000	0.000	0.0	0.0
12	0.000	0.000	0.0	0.0	12	0.000	0.000	0.0	0.0
13	0.849	0.849	1239.7	13.2	13	0.887	0.887	1463.0	8.5
14	0.938	0.938	1297.7	10.7	14	0.921	0.921	1363.9	8.9
15	0.797	0.797	233.8	6.0	15	0.747	0.748	83.6	3.7
16	0.933	0.933	957.2	3.9	16	0.984	0.984	822.6	2.3
17	0.538	0.538	92.8	3.6	17	0.653	0.651	35.7	3.9
18	0.860	0.860	826.7	9.6	18	0.000	0.000	0.0	0.0
19	0.000	0.000	0.0	0.0	19	0.889	0.920	1.4	0.3
20	0.924	0.924	1955.2	7.3	20	0.948	0.948	2945.3	9.1
21	0.954	0.954	377.2	3.4	21	0.975	0.975	504.4	3.3
22	0.945	0.945	1442.4	6.7	22	0.969	0.969	1697.3	4.9
23	0.000	0.000	0.0	0.0	23	0.000	0.000	0.0	0.0
24	0.000	0.000	0.0	0.0	24	0.000	0.000	0.0	0.0
25	0.000	0.000	0.0	0.0	25	0.000	0.000	0.5	0.7
26	0.000	0.000	0.0	0.0	26	0.000	0.000	0.0	0.0
27	0.000	0.000	0.0	0.0	27	0.000	0.000	0.0	0.0
28	0.000	0.000	0.0	0.0	28	0.000	0.000	0.0	0.0
29	0.000	0.000	0.0	0.0	29	0.000	0.000	0.0	0.0
30	0.000	0.000	0.0	0.0	30	0.000	0.000	0.0	0.0

investigated. For each target region the number of connections from voxels in the thalamus to the target region was found. This information was output as an image for each target region. The outputs produced for each target region were then combined using "find_the_biggest" (http://www.fmrib.ox.ac.uk/fsl/fdt/fdt_biggest.html). This process involved classifying each voxel in the thalamus depending on the target to which the greatest number of the paths from that voxel propagated. This process was a "hard segmentation" in which a single label was assigned to each thalamic voxel.

Reproducibility

Single dataset processed multiple times

For the initial assessment of reproducibility a single dataset was processed 16 times. The implementation of the method in FMRIB's Diffusion Toolbox (FDT v2.0, http://www.fmrib.ox.ac.uk/fsl/fdt/index.html) is by default set up with a specified seed for the random number generator (12345). The seed value used in the probabilistic tractography algorithm determines the sequence of random numbers generated. To allow some variation in the output results, the seed was changed each time the dataset was processed. Fifteen additional seed values were chosen for processing (23456, 34567, 45678, 56789, 67891, 78912, 89123, 15432, 26543, 37654, 48765, 59876, 61987, 72198, and 83219). Probabilistic tractography was carried out on the same dataset 16 times using the default seed and the 15 other values specified. For each set of results a hard segmentation was carried out to produce 16 segmentations of the thalamus. From each of these 16 resulting segmentations, the volume of the thalamic region assigned to each target region was determined and the mean and standard deviation were calculated. To determine the reproducibility of the segmentation, the hard segmentation results were then compared using the overlap calculation described below. The analysis was first carried out using the six large cortical regions (shown in Table 2) and then using the 31 smaller, cortical target regions (Table 1).

Overlap

The overlap between each pair of segmentation results and the total accumulated overlap for all segmentation results were investigated using the method described by Crum et al. (2006). The calculation of the overlap values was based on the Tanimoto Coefficient (TC). For two regions, A and B, the TC is given by:

$$\mathrm{TC} = \frac{N(A \cap B)}{N(A \cup B)}$$

To compare the overlap between a pair of images, each containing n regions, TC was extended as follows to give the Overlap By Pairs (OBP) value (Crum et al., 2006):

$$\mathrm{OBP} = \frac{\sum_{i=1:n} \alpha_i N(A_i \cap B_i)}{\sum_{i=1:n} \alpha_i N(A_i \cup B_i)}$$

where i is the region label and α_i is the weighting. Two weightings were used: $\alpha_i = 1$ and $\alpha_i = 1/V_i$, where:

$$V_i = \frac{N(A_i) + N(B_i)}{2}$$

When $\alpha_i = 1$, corresponding to the standard Jaccard index, large regions contributed more to the total overlap, approximately in

Table 6
Fuzzy overlap OBL values for regions classified as connecting to each of the 31 cortical target regions (single subject, 16 repetitions of probabilistic tractography analysis).

Region	Left thalamus		Region	Right thalamus	
	Fuzzy overlap (OBL)			Fuzzy overlap (OBL)	
	$\alpha_i=1$	$\alpha_i=1/V_i$		$\alpha_i=1$	$\alpha_i=1/V_i$
1	0.681	0.681	1	0.767	0.767
2	0.770	0.770	2	0.791	0.791
3	0.753	0.753	3	0.825	0.825
4	0.825	0.825	4	0.490	0.490
5	0.744	0.744	5	0.493	0.493
6	0.766	0.766	6	0.845	0.845
7	0.872	0.872	7	0.046	0.046
8	0.022	0.022	8	0.677	0.677
9	0.740	0.740	9	0.791	0.791
10	0.852	0.852	10	0.718	0.718
11	0.000	0.000	11	0.002	0.003
12	0.000	0.000	12	0.000	0.000
13	0.893	0.893	13	0.900	0.900
14	0.925	0.925	14	0.923	0.923
15	0.835	0.835	15	0.823	0.823
16	0.936	0.936	16	0.976	0.976
17	0.642	0.642	17	0.643	0.643
18	0.876	0.876	18	0.121	0.121
19	0.473	0.473	19	0.703	0.703
20	0.939	0.939	20	0.942	0.942
21	0.924	0.924	21	0.962	0.962
22	0.935	0.935	22	0.961	0.961
23	0.388	0.388	23	0.431	0.431
24	0.003	0.003	24	0.310	0.310
25	0.577	0.577	25	0.332	0.333
26	0.016	0.019	26	0.245	0.245
27	0.166	0.166	27	0.174	0.173
28	0.000	0.000	28	0.000	0.000
29	0.037	0.038	29	0.403	0.403
30	0.000	0.000	30	0.000	0.000
31	0.645	0.645	31	0.547	0.547

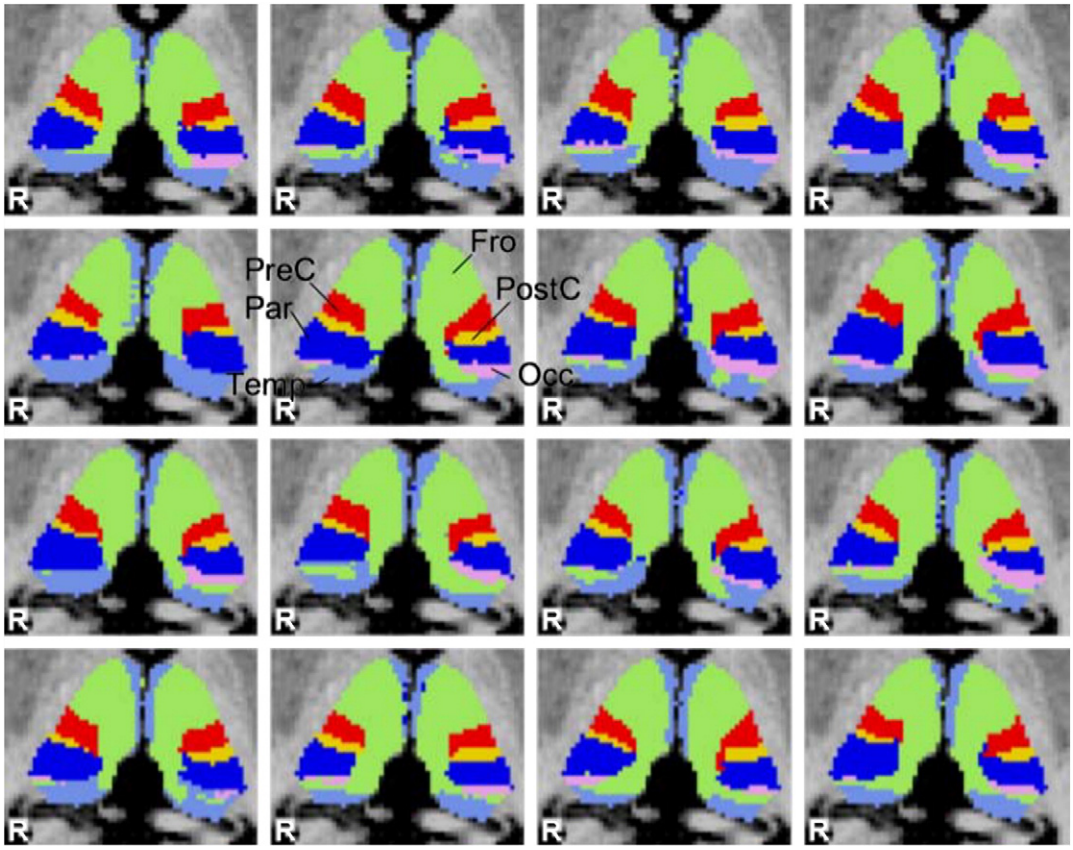

Fig. 4. Results of the hard segmentation of the thalamus using six cortical target regions. 16 datasets were acquired from the same subject. For each of the 16 datasets from the same subject, a single axial slice of the segmentation results is shown. This slice corresponds to Slice 8 in Fig. 2.

proportion to their volume. When $\alpha_i = 1/V_i$ all regions contributed equally to the overall overlap regardless of size. Using this second weighting meant that misalignment of smaller regions would be more apparent in OBP and TAO values.

For a group including m pairs of images, the "groupwise" overlap measure (Total Accumulated Overlap, TAO) was determined:

$$\mathrm{TAO} = \frac{\sum_m \sum_{i=1:n} \alpha_i N(A_i \cap B_i)}{\sum_m \sum_{i=1:n} \alpha_i N(A_i \cup B_i)}$$

Finally, overlap-by-label (OBL) values were calculated, for which overlap of each region across all datasets was calculated:

$$\mathrm{OBL} = \frac{\sum_m \alpha_i N(A_i \cap B_i)}{\sum_m \alpha_i N(A_i \cup B_i)}$$

In addition to calculating overlap of hard segmentation results, overlap of results output from "probtrackx" showing number of paths to each cortical target region was also investigated. Initially, the proportion of paths from each voxel to each target region was determined. At each voxel, this was calculated by dividing the number of paths to a particular target region by the total number of paths to all target regions: all voxels had a value between 0 and 1. Fuzzy overlap values were then calculated using similar equations to those shown above, with $N\ (A_i \cap B_i)$ replaced by $\sum_{\mathrm{voxels\,v}} \mathrm{MIN}(A_\mathrm{v}, B_\mathrm{v})$ and $N\ (A_i \cup B_i)$ replaced with $\sum_{\mathrm{voxels\,v}} \mathrm{MAX}(A_\mathrm{v}, B_\mathrm{v})$ (Crum et al., 2006). For example, the equation for fuzzy overlap TAO is:

$$\text{Fuzzy Overlap TAO} = \frac{\sum_m \sum_{i=1:n} \left(\alpha_i \sum_{\mathrm{voxels\,v}} \mathrm{MIN}(A_\mathrm{v}, B_\mathrm{v}) \right)}{\sum_m \sum_{i=1:n} \left(\alpha_i \sum_{\mathrm{voxels\,v}} \mathrm{MAX}(A_\mathrm{v}, B_\mathrm{v}) \right)}$$

Table 7
OBL values and average volumes of regions classified as connecting to each of the six cortical target regions (16 datasets from a single subject).

Region	Left thalamus				Region	Right thalamus			
	Overlap (OBL)		Volume (mm^3)			Overlap (OBL)		Volume (mm^3)	
	$\alpha_i = 1$	$\alpha_i = 1/V_i$	Mean	sd		$\alpha_i = 1$	$\alpha_i = 1/V_i$	Mean	sd
Frontal	0.707	0.706	3957.4	322.7	Frontal	0.774	0.774	4202.0	243.9
Occipital	0.506	0.504	447.3	86.5	Occipital	0.231	0.191	121.7	89.0
Parietal	0.621	0.621	1534.3	239.8	Parietal	0.716	0.715	1977.6	173.9
Post-Central	0.507	0.498	358.9	82.6	Post-Central	0.377	0.364	318.9	103.9
Pre-Central	0.680	0.680	969.7	116.4	Pre-Central	0.699	0.700	865.6	82.5
Temporal	0.649	0.645	2457.3	290.8	Temporal	0.701	0.701	2233.0	183.3

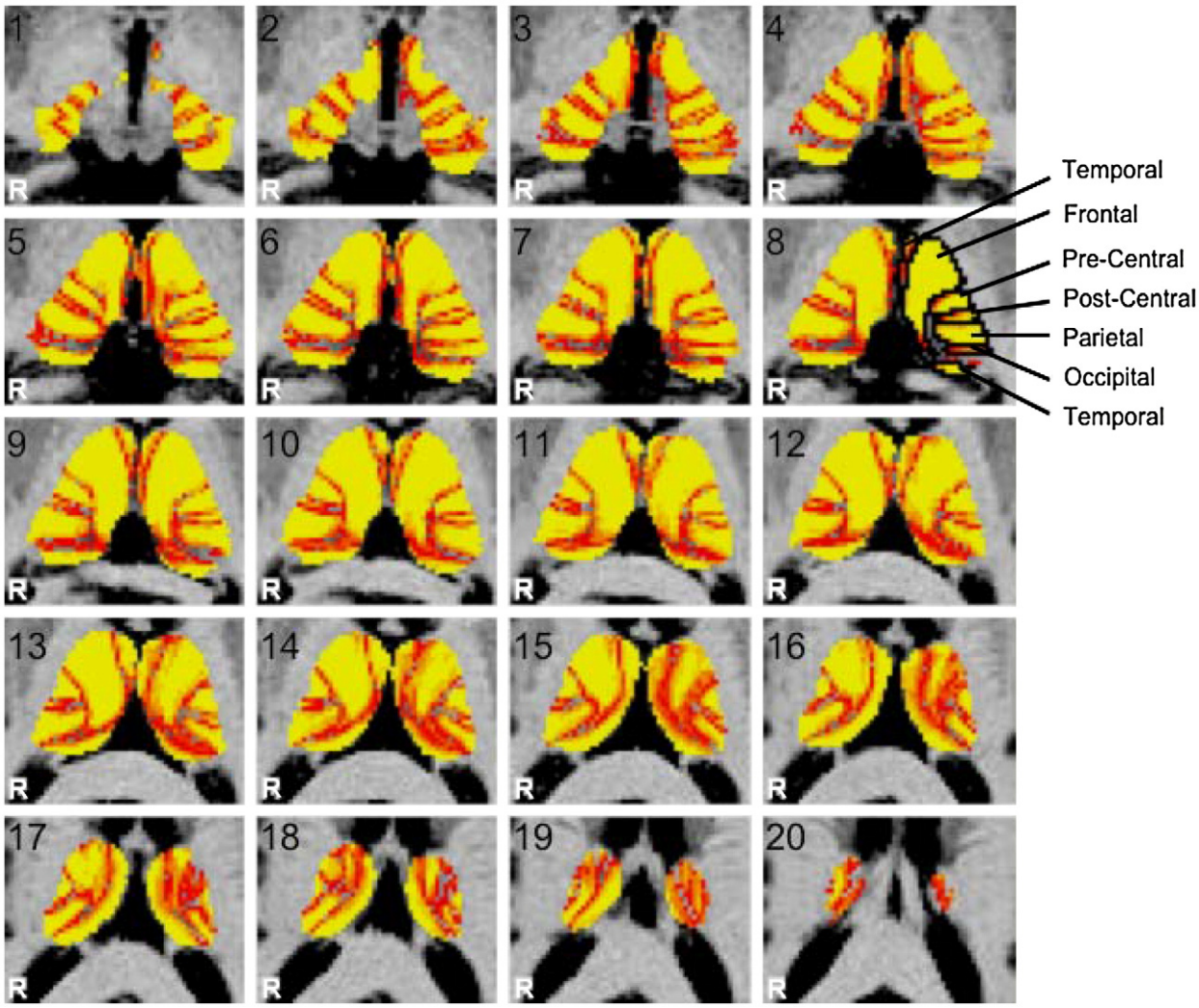

Fig. 5. The overlap of hard segmentation results from 16 datasets from the same subject. The number of datasets in which thalamic voxels had the same region classification when six cortical target regions were used is shown, from 9 datasets (red) to 16 datasets (yellow). 20 axial slices through the thalami are shown, from inferior (Slice 1) to superior (Slice 20).

Multiple datasets from a single subject

Probabilistic tractography was then carried out using 16 different DTI datasets from the same subject (using the default seed value in each case). Fuzzy overlap values were calculated based on these results. As for processing of the single dataset, a hard segmentation of the results from probtrackx was carried out to produce 16 thalamic

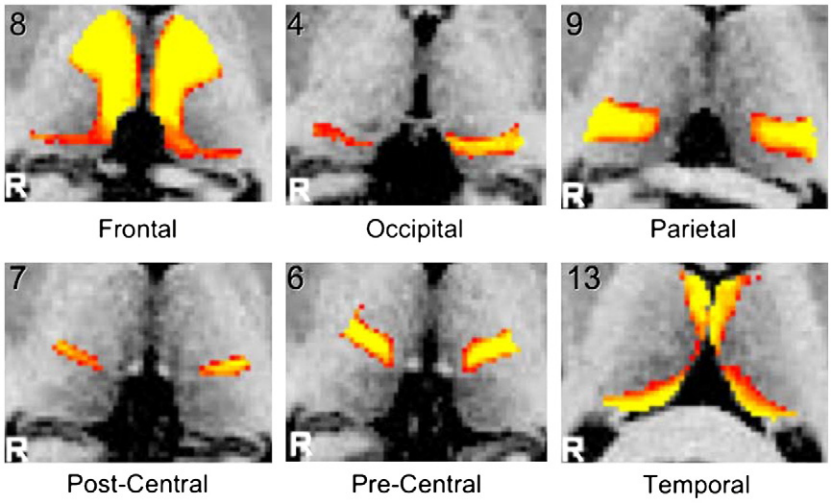

Fig. 6. Example slice for each target region showing overlap of the thalamic sub-region in at least 5 of the 16 datasets for a single subject. Numbers in the top left corner of each image correspond to the slice number from Fig. 5.

Table 8
Fuzzy overlap OBL values for regions classified as connecting to each of the six cortical target regions (16 datasets from a single subject).

Region	Left thalamus		Region	Right thalamus	
	Fuzzy overlap (OBL)			Fuzzy overlap (OBL)	
	$\alpha_i = 1$	$\alpha_i = 1/V_i$		$\alpha_i = 1$	$\alpha_i = 1/V_i$
Frontal	0.721	0.721	Frontal	0.780	0.780
Occipital	0.510	0.511	Occipital	0.310	0.287
Parietal	0.634	0.634	Parietal	0.705	0.705
Post-central	0.511	0.503	Post-central	0.420	0.411
Pre-central	0.682	0.682	Pre-central	0.701	0.702
Temporal	0.659	0.658	Temporal	0.710	0.709

segmentations. As the same structural image was used with all 16 DTI datasets, no additional registration of image data was required because segmentations were output in the same structural space. The volume of the thalamic region assigned to each target region was determined and the mean and standard deviations for each target region were calculated. The overlap values were found for these different datasets using the formulae for OBP, OBL and TAO shown above. A map was also generated showing the voxels that were classified as belonging to the same region in more than half of the 16 datasets. As for the analysis of a single dataset processed 16 times, the processing described above was first carried out using the six large cortical regions (Table 2) and then using the 31 smaller, cortical target regions (Table 1).

Multiple subjects

Next, connectivity-based segmentation of the thalamus was carried out in 18 control datasets. It was expected that, for analysis of single subject data, there would be limited overlap of many of the regions identified using 31 target regions; hence, for the normal subjects, analysis was only carried out using the six larger cortical target regions. Cortical and sub-cortical regions were found for each subject. Based on the preliminary results of the single subject analyses (see below) the segmentations were performed only on the inverted-contrast T1 maps, not the IR-FSPGR images.

For these subjects, in addition to generating matrices describing the transformation from diffusion to structural space, matrices describing the transformation from structural to a standard space were also required. All structural images were normalised to the Montreal Neurological Institute (MNI) 152 standard T1-weighted MRI template (with 1 mm isotropic voxels). Spatial normalisation was achieved for each subject using FNIRT (Andersson et al., 2007), with the inverted-contrast T1 map used as the input image and the 1 mm-resolution T1-weighted MNI space image as the reference image. The default configuration file provided with FSL (http://www.fmrib.ox.ac.uk/fsl/fnirt/index.html#fnirt_config_files) was used.

For each subject, the size of each of the regions identified by the segmentation was normalised. First, the total intra-cranial volume (TIV) was calculated as the total volume of the BET mask which included GM, WM and CSF. The average TIV for the group of normal volunteers was then calculated. Region sizes were normalised by dividing by the subjects TIV and multiplying by the group average TIV. Average normalised volumes across the group were determined. Overlap of the segmentation results from all normal subjects was investigated by calculating OBP, OBL and TAO values.

For calculation of fuzzy overlap values in the group of normal subjects, results from probtrackx, prior to hard segmentation, were non-linearly transformed into standard space. At each voxel, the proportion of paths to each target region was determined, and fuzzy overlap values were calculated based on these results.

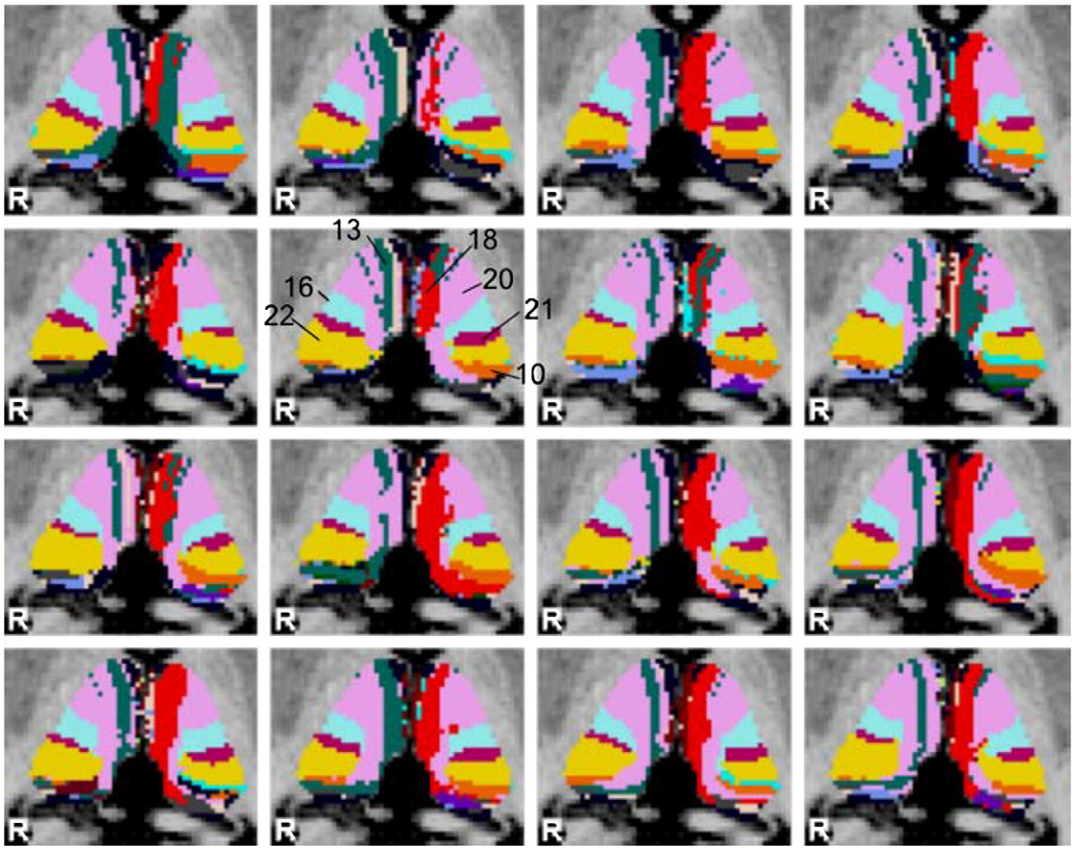

Fig. 7. Results of the hard segmentation of the thalamus using 31 cortical target regions. For each of the 16 datasets from the same subject, a single axial slice of the segmentation results is shown. This slice corresponds to Slice 8 in Fig. 3.

Table 9
OBL values and average volumes of regions classified as connecting to each of the 31 cortical target regions (16 datasets from a single subject).

Region	Left thalamus				Region	Right thalamus			
	Overlap (OBL)		Volume (mm^3)			Overlap (OBL)		Volume (mm^3)	
	$\alpha_i=1$	$\alpha_i=1/V_i$	Mean	sd		$\alpha_i=1$	$\alpha_i=1/V_i$	Mean	sd
1	0.054	0.034	7.6	11.6	1	0.031	0.023	7.6	15.9
2	0.263	0.247	321.9	185.7	2	0.199	0.186	212.5	99.4
3	0.159	0.168	116.5	78.2	3	0.130	0.107	230.3	227.7
4	0.090	0.057	95.1	131.3	4	0.016	0.006	24.1	58.2
5	0.144	0.139	30.5	21.6	5	0.024	0.025	7.7	15.6
6	0.112	0.070	122.3	119.0	6	0.134	0.110	137.0	107.2
7	0.120	0.092	126.7	89.9	7	0.000	0.000	2.4	9.6
8	0.000	0.000	0.0	0.0	8	0.000	0.000	11.0	43.6
9	0.091	0.092	184.0	148.4	9	0.109	0.107	295.6	166.4
10	0.522	0.522	539.6	109.7	10	0.259	0.216	146.8	91.2
11	0.000	0.000	0.0	0.0	11	0.000	0.000	0.0	0.0
12	0.000	0.000	0.0	0.0	12	0.000	0.000	0.0	0.0
13	0.096	0.071	405.4	457.8	13	0.407	0.404	1198.9	278.8
14	0.510	0.507	1417.2	327.6	14	0.548	0.536	1311.8	304.1
15	0.095	0.084	161.6	141.0	15	0.014	0.009	13.2	24.1
16	0.669	0.669	997.5	119.3	16	0.697	0.698	867.5	83.3
17	0.025	0.014	13.2	24.2	17	0.102	0.070	44.2	40.6
18	0.362	0.340	1108.4	465.5	18	0.000	0.000	0.2	0.5
19	0.000	0.000	0.3	1.0	19	0.011	0.004	1.6	4.3
20	0.604	0.604	2349.1	188.4	20	0.623	0.621	2832.6	309.6
21	0.508	0.499	360.0	81.5	21	0.378	0.366	319.8	104.0
22	0.624	0.623	1359.8	206.0	22	0.711	0.710	2013.1	184.5
23	0.000	0.000	0.2	0.5	23	0.000	0.000	0.2	0.5
24	0.000	0.000	0.0	0.0	24	0.000	0.000	0.1	0.3
25	0.000	0.000	6.1	22.2	25	0.000	0.000	0.1	0.3
26	0.000	0.000	0.0	0.0	26	0.000	0.000	0.0	0.0
27	0.000	0.000	0.0	0.0	27	0.000	0.000	0.0	0.0
28	0.000	0.000	0.0	0.0	28	0.000	0.000	0.0	0.0
29	0.000	0.000	0.0	0.0	29	0.008	0.002	2.2	6.6
30	0.000	0.000	0.0	0.0	30	0.000	0.000	0.0	0.0
31	0.003	0.005	1.7	5.1	31	0.055	0.035	38.3	58.3

Results

Cortical and sub-cortical regions specific to the first subject were generated based on the IR-FSPGR image and the inverted-contrast T1 map. Fig. 1 shows an axial slice from each of these structural images, along with the corresponding cortical and sub-cortical segmentations derived from these images. Cortical regions defined from both structural images were similar; TAO between the two sets of results was 0.753. Differences were observed in the thalamus masks generated from the two images: for the left thalamus OBP was 0.629, for the right thalamus OBP was 0.557. Comparing the results to the probabilistic map shown in Ahsan et al. (2007), the lateral edge of both thalami was excluded in the masks generated based on the IR-FSPGR image (identified by the arrows in Fig. 1). The masks obtained using the inverted-contrast T1 map were more inclusive, resulting in thalami with a similar shape to those shown in Ahsan et al. It was noted that the contrast between the thalamus and the adjacent white matter (internal capsule) was much greater in the inverted-contrast T1 map than in the IR-FSPGR image. For this reason, cortical and thalamic regions found using the inverted-contrast T1 map were used in the probabilistic tractography processing.

Single dataset processed multiple times

Six cortical target regions

Fig. 2 shows an example of the segmentation obtained using the 6 large, cortical target regions. Both right and left thalami were found to have connections to all six target regions. Comparable patterns of connectivity were seen in the right and left thalami, and for both thalami the largest region was classified as connecting to the frontal cortical target region. OBP values were calculated for all 120 possible pairs of images. For the right thalamus, OBP values found using $\alpha_i=1$ varied between 0.960 and 0.972, with a mean and standard deviation of 0.966 and 0.003, respectively. Using $\alpha_i=1$, OBP values for the left thalamus ranged from 0.940 to 0.956, with a mean value of 0.949 and a standard deviation of 0.003. Using $\alpha_i=1/V_i$, OBP values for the right thalamus varied between 0.877 and 0.940, and for the left thalamus between 0.919 and 0.945. In addition to OBP values, TAO values were found for right and left thalami as a measure of the overlap between all 16 thalamic segmentations. For the right thalamus, TAO values were 0.966 for $\alpha_i=1$ and 0.908 for $\alpha_i=1/V_i$. TAO values for the left thalamus were 0.949 for $\alpha_i=1$ and 0.934 for $\alpha_i=1/V_i$. OBL values for all regions are shown in Table 3. Fuzzy overlap OBL values are shown in Table 4; fuzzy overlap TAO values for the right thalamus were 0.960 ($\alpha_i=1$) and 0.919 ($\alpha_i=1/V_i$) and for the left thalamus were 0.941 ($\alpha_i=1$) and 0.924 ($\alpha_i=1/V_i$).

For each of the 16 results, the volume of the regions labelled as being connected to each target region was determined. The mean and standard deviation of these results are shown, together with the OBL values, in Table 3. For both left and right thalami the region classified as being connected to the frontal region had the largest volume. The number of results in which each voxel was classified by a particular label was also determined. Most voxels in the thalamus were labelled as connecting to the same region each time the dataset was processed. 94.8% of voxels in the right thalamus and 92.2% of voxels in the left thalamus were given the same label in all 16 results.

31 Cortical target regions

Reproducibility was also investigated using the 31 cortical target regions shown in Table 1. Fig. 3 shows an example of the results produced using these target regions. OBP values for all pairs of images and TAO values were calculated for both $\alpha_i=1$ and $\alpha_i=1/V_i$. For $\alpha_i=1$, OBP values were greater than 0.879 for all

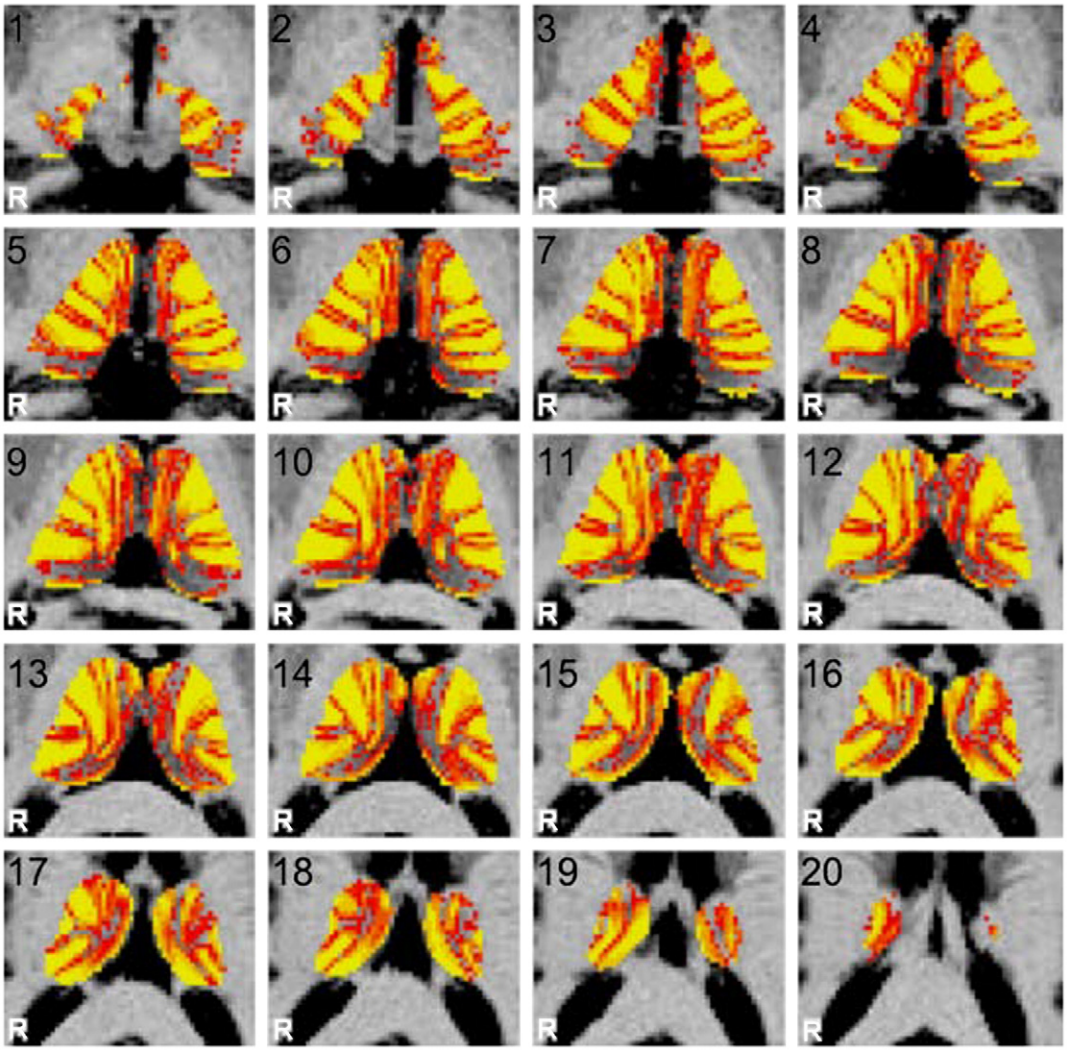

Fig. 8. The overlap of hard segmentation results from 16 datasets from the same subject. The number of datasets in which thalamic voxels had the same region classification when 31 cortical target regions were used, from 9 datasets (red) to 16 datasets (yellow). 20 axial slices through the thalami are shown, from inferior (Slice 1) to superior (Slice 20).

pairs of images. OBP values for $\alpha_i = 1/V_i$ were lower and showed more variation across the dataset, ranging from 0.631 to 0.860 for the right thalamus and from 0.597 to 0.755 for the left thalamus. TAO values for the right thalamus were 0.921 ($\alpha_i = 1$) and 0.721 ($\alpha_i = 1/V_i$), and for the left thalamus were 0.890 ($\alpha_i = 1$) and 0.675 ($\alpha_i = 1/V_i$). Table 5 shows OBL values for all regions in the left and right thalami.

For each of the 16 results, the volume labelled as being connected to each target region was determined. The average values are also shown in Table 5. The results showed that there were large regions of the thalamus that were connected to Regions 13 (frontal lobe), 14 (posterior temporal lobe), 16 (pre-central gyrus), 20 (superior frontal gyrus) and 22 (superior parietal gyrus). In addition, from the left thalamus, but not right, there was a large region classified as connecting to Region 18 (anterior orbital gyrus). It is also notable that there were several regions to which no thalamic connections were found, especially from right thalamus. In processing this dataset 16 times, the majority of voxels in the thalamus were found to have the same classification in all 16 results (88.6% for right thalamus and 83.8% for left thalamus).

Using 31 cortical target regions, fuzzy overlap OBL values obtained are shown in Table 6. For the right thalamus, fuzzy overlap TAO values were 0.913 ($\alpha_i = 1$) and 0.532 ($\alpha_i = 1/V_i$); for the left thalamus, fuzzy overlap TAO values were 0.891 ($\alpha_i = 1$) and 0.554 ($\alpha_i = 1/V_i$).

Multiple datasets from the same subject

Following the assessment of reproducibility of the thalamic segmentation method for a single dataset, a similar assessment was carried out using 16 datasets from the same subject. Processing was carried out using six and 31 cortical target regions.

Six cortical target regions

In Fig. 4, a single axial slice of the thalamic segmentation is shown for each of the 16 datasets. As all the datasets used the same structural image, and all results were output in structural space, no transformation between datasets was required. Fig. 4 shows that the general pattern of connectivity was similar in all results; however, there was some variability between the datasets. TAO values for the right thalamus were 0.713 for $\alpha_i = 1$ and 0.540 for $\alpha_i = 1/V_i$, and for the left thalamus were 0.657 for $\alpha_i = 1$ and 0.605 for $\alpha_i = 1/V_i$. OBL values are shown in Table 7. Fig. 5 shows voxels that had the same label in more than half of the 16 results. 55.5% of voxels in the right thalamus and 44.2% of voxels in the left thalamus were classified by the same label in all 16 results. For each thalamic region identified following hard segmentation, voxels which had the same label in at least 5 datasets were identified. Fig. 6 shows examples of the results obtained for each target region, showing that the thalamic regions are very similar even when the overlap "threshold" is reduced from 9 (Fig. 5) to 5 (Fig. 6).

Table 10
Fuzzy overlap OBL values for regions classified as connecting to each of the 31 cortical target regions (16 datasets from a single subject).

Region	Left thalamus		Region	Right thalamus	
	Fuzzy overlap (OBL)			Fuzzy overlap (OBL)	
	$\alpha_i=1$	$\alpha_i=1/V_i$		$\alpha_i=1$	$\alpha_i=1/V_i$
1	0.212	0.211	1	0.191	0.184
2	0.324	0.321	2	0.294	0.288
3	0.312	0.314	3	0.292	0.306
4	0.203	0.189	4	0.173	0.175
5	0.256	0.257	5	0.136	0.137
6	0.250	0.231	6	0.240	0.220
7	0.260	0.237	7	0.053	0.040
8	0.025	0.024	8	0.167	0.156
9	0.320	0.317	9	0.342	0.339
10	0.507	0.507	10	0.303	0.280
11	0.000	0.000	11	0.004	0.003
12	0.000	0.000	12	0.000	0.000
13	0.363	0.365	13	0.503	0.501
14	0.521	0.519	14	0.550	0.541
15	0.304	0.307	15	0.152	0.135
16	0.682	0.682	16	0.701	0.702
17	0.058	0.043	17	0.174	0.151
18	0.437	0.431	18	0.028	0.026
19	0.124	0.121	19	0.138	0.136
20	0.628	0.629	20	0.662	0.661
21	0.512	0.504	21	0.419	0.411
22	0.635	0.635	22	0.701	0.700
23	0.078	0.076	23	0.091	0.100
24	0.000	0.000	24	0.130	0.105
25	0.125	0.123	25	0.055	0.045
26	0.007	0.006	26	0.095	0.101
27	0.028	0.019	27	0.010	0.004
28	0.000	0.000	28	0.000	0.000
29	0.011	0.007	29	0.085	0.071
30	0.000	0.000	30	0.000	0.000
31	0.135	0.131	31	0.219	0.199

Table 7 also shows the mean volumes of the thalamic regions classified as being connected to each cortical target region across the 16 acquisitions, determined from results of hard segmentation.

Fuzzy overlap TAO values were 0.717 ($\alpha_i=1$) and 0.576 ($\alpha_i=1/V_i$) for the right thalamus and 0.667 ($\alpha_i=1$) and 0.614 ($\alpha_i=1/V_i$) for the left thalamus. Fuzzy overlap OBL values are shown in Table 8.

31 cortical target regions

Reproducibility was also assessed using the 16 datasets and the 31 cortical target regions (Table 1). Results of the hard segmentation of the thalamus, using 31 target regions, are shown in Fig. 7. A single axial slice is shown for each of the 16 datasets. Using $\alpha_i=1$, OBP values varied between 0.404 and 0.631 for the right thalamus and between 0.341 and 0.596 for the left thalamus. Using $\alpha_i=1/V_i$, a maximum OBP of 0.326 was obtained. The TAO values for the right thalamus were 0.520 ($\alpha_i=1$) and 0.189 ($\alpha_i=1/V_i$), and for the left thalamus were 0.462 ($\alpha_i=1$) and 0.215 ($\alpha_i=1/V_i$); Table 9 shows corresponding OBL values.

The mean volume classified as being connected to each target region in the hard segmentation results is also shown in Table 9. As seen for the single dataset, large regions of the thalamus were identified as being connected to Regions 14 (posterior temporal lobe), 16 (pre-central gyrus), 20 (superior frontal gyrus) and 22 (superior parietal gyrus), from both thalami and 18 (anterior orbital gyrus) from the left thalamus. Fig. 8 shows the voxels that were classified as the same value in more than 50% of the 16 results. The results appeared to be most consistent along the lateral edges of the left and right thalami, with more variation between the results seen in the medial and posterior regions. 30.7% of voxels in the right thalamus and 25.4% of voxels in the left thalamus were given the same classification in the results from all 16 datasets.

Fuzzy overlap OBL values obtained when 31 target regions were used are shown in Table 10. For the right thalamus, fuzzy overlap TAO values were 0.551 ($\alpha_i=1$) and 0.204 ($\alpha_i=1/V_i$); for the left thalamus, fuzzy overlap TAO values were 0.501 ($\alpha_i=1$) and 0.236 ($\alpha_i=1/V_i$).

Multiple subjects

Six cortical target regions

Fig. 9 shows a single axial slice of the thalamus segmentation for each of the 18 subjects when six cortical target regions were used. The results are shown in standard space and overlaid on a T1-weighted MNI space image. Using $\alpha_i=1$, the OBP values varied between 0.310 and 0.533 for the right thalamus and between 0.250 and 0.561 for the left thalamus. For $\alpha_i=1/V_i$, OBP values in the right thalamus were between 0.158 and 0.426, and between 0.153 and 0.544 for the left thalamus. TAO values were 0.430 ($\alpha_i=1$) and 0.280 ($\alpha_i=1/V_i$) for the right thalamus and 0.407 ($\alpha_i=1$) and 0.287 ($\alpha_i=1/V_i$) for the left thalamus. OBL values for this group of normal subjects are shown in Table 11.

Table 11 also shows the mean adjusted size of each of the regions from the segmentation results from all subjects. The pattern of region sizes was similar to that found for the single subject, with the largest volume classified as connecting to the frontal target region. As expected, the standard deviations for these results were much greater than those obtained for the single subject.

Fig. 10 shows the voxels that were classified as connected to the same cortical region in more than half of the 18 results. Classification of voxels is similar for the regions identified as being connected to the frontal, parietal, temporal and pre-central target regions. Rather less overlap was seen in the regions classified as being connected to the post-central and occipital cortical targets. In Fig. 11, voxels classified by the same label in at least 5 of the 18 normal subjects are shown; one axial slice is displayed for each of the six cortical target regions, showing that the thalamic regions are very similar even when the overlap "threshold" is reduced from 10 (Fig. 10) to 5 (Fig. 11).

Fuzzy overlap OBL values calculated from normal subject data are shown in Table 12. Fuzzy overlap TAO values were 0.469 ($\alpha_i=1$) and 0.336 ($\alpha_i=1/V_i$) for the right thalamus and 0.469 ($\alpha_i=1$) and 0.330 ($\alpha_i=1/V_i$) for the left thalamus.

TAO values for all the analyses of single and multiple subject data are summarised in Table 13; fuzzy overlap TAO values are summarised in Table 14.

Discussion

This work demonstrates that using connectivity-based analysis to segment the thalamus on MR images of the human brain yields subdivisions with high reproducibility. Although the method relies on repeated sampling of a Probability Density Function (PDF), we showed that segmentations produced by reanalysing the same dataset using different starting conditions were extremely similar. Additionally, we showed within-subject reliability of the method with stable segmentation results obtained from multiple imaging acquisitions from the same subject. For this type of method, measurement of the within-subject reliability can be used as an estimate of the upper bound on the level of inter-subject reliability. Finally we were able to show similar thalamic segmentations across a group of 18 normal subjects, preserving relative sizes and positions of thalamic segments. This method had previously shown great promise; here we demonstrate a very high level of reliability, which will underpin future implementations of this method to examine thalamic subdivisions in health and disease.

The connectivity-based segmentation method implemented uses data obtained from tractography, a method of tracing paths through

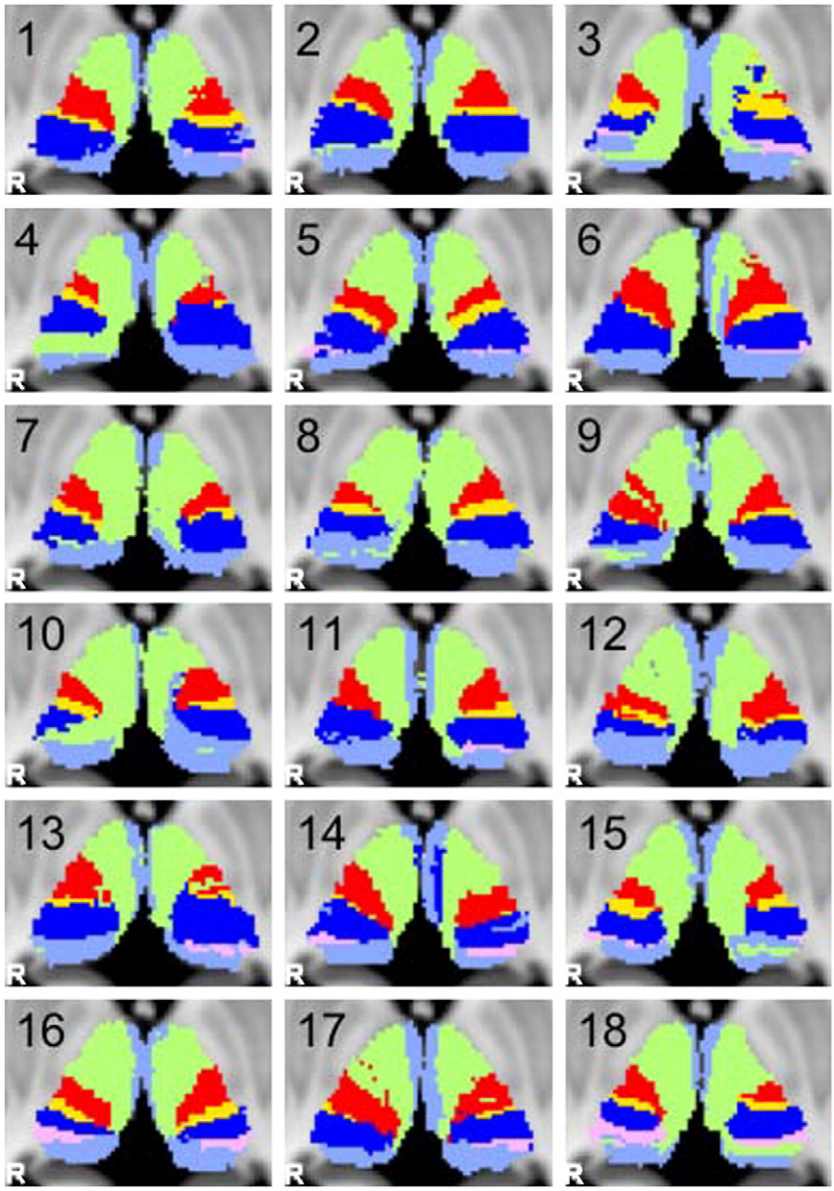

Fig. 9. Results of thalamic segmentation in 18 subjects using six cortical target regions. A single axial slice in standard space is shown for each subject overlaid on the T1-weighted MNI space image available in FSL (www.fmrib.ox.ac.uk/fsl/).

diffusion imaging data. There are two main tractography techniques: deterministic tractography and probabilistic tractography. Strengths and weaknesses of tractography methods have been debated (Ciccarelli et al., 2008). Probabilistic tractography, as implemented by Behrens et al. (2003a,b), has been used to examine connectivity between target and seed regions based on likelihood of connection. The pathways of connection are not themselves necessarily the phenomena of interest, nor are all the possible pathways identified by repeated PDF sampling necessarily likely on the basis of known anatomical connections. Here, we have examined only the

Table 11
OBL values and average volumes of regions from the 18 normal datasets assigned to each of the six target regions. For each subject, the region sizes were adjusted for differences in the total intracranial volume of subjects before the mean for all subjects was calculated.

Region	Left thalamus				Region	Right thalamus			
	Overlap (OBL)		Volume (mm^3)			Overlap (OBL)		Volume (mm^3)	
	$\alpha_i = 1$	$\alpha_i = 1/V_i$	Mean	sd		$\alpha_i = 1$	$\alpha_i = 1/V_i$	Mean	sd
Frontal	0.480	0.468	2673.8	611.6	Frontal	0.550	0.552	3163.3	436.5
Occipital	0.124	0.079	196.3	176.9	Occipital	0.060	0.030	120.4	186.8
Parietal	0.450	0.440	1775.1	507.7	Parietal	0.393	0.383	1448.0	422.1
Post-Central	0.148	0.129	356.2	207.4	Post-central	0.128	0.108	257.6	128.8
Pre-Central	0.314	0.302	790.3	340.3	Pre-central	0.335	0.334	861.9	334.2
Temporal	0.413	0.412	2467.5	519.7	Temporal	0.420	0.419	2189.9	381.7

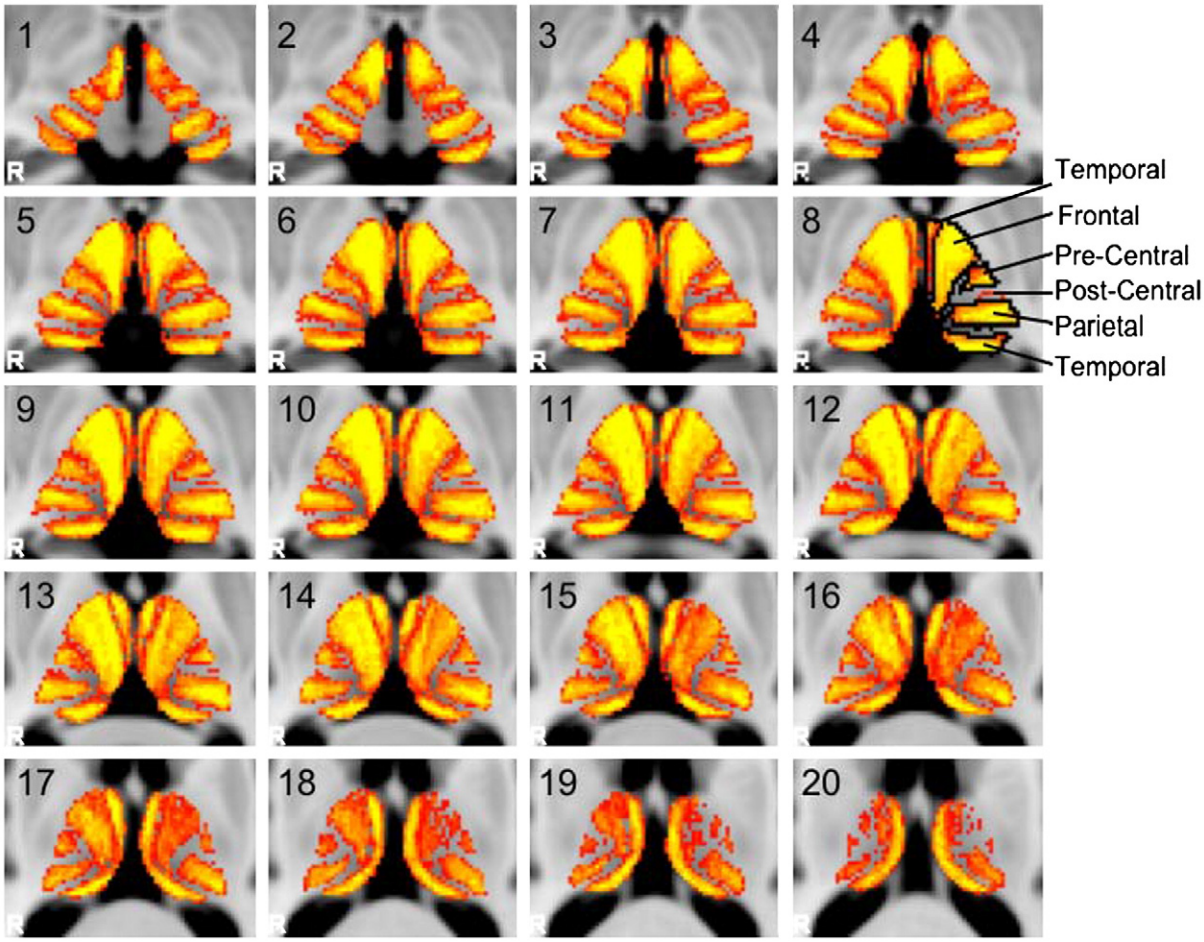

Fig. 10. The overlap of hard segmentation results from 18 subjects. The number of datasets in which thalamic voxels had the same region classification when six cortical target regions were used, from 10 datasets (red) to 18 datasets (yellow). 20 axial slices through the thalami are shown in standard space, from inferior (Slice 1) to superior (Slice 20).

reproducibility of thalamic segmentation between subjects using this technique; we have not yet explored the thalamo-cortical connections identified using this method, or their relationship with known or hypothesised thalamo-cortical and cortico-thalamic connections.

In Behrens et al.'s connectivity-based segmentation method (Behrens et al., 2003a,b), 10,000 paths were generated from each voxel contained within a thalamus seed mask. Each voxel in the thalamus was then classified depending on the target region in the cortex to which it had the greatest number of connections. This connectivity-based segmentation produced similar results for both the left and right thalami. The researchers proposed that the regions identified corresponded to thalamic nuclei or nuclear groups. The

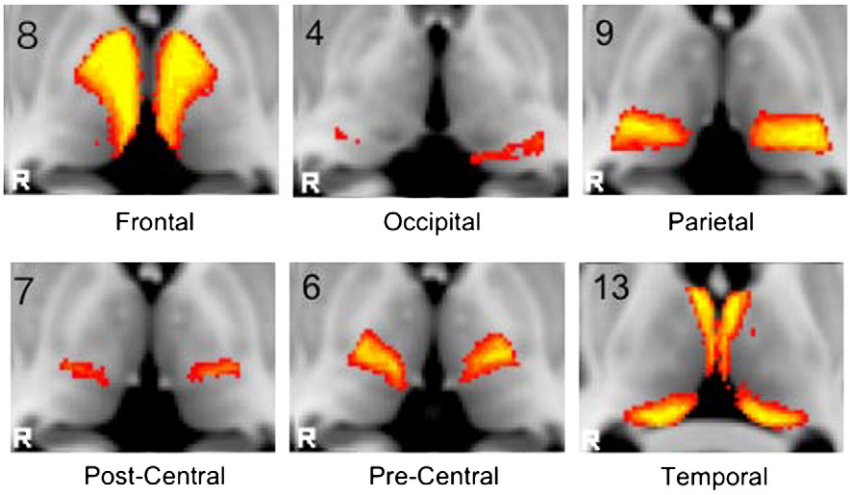

Fig. 11. Example slice for each target region showing overlap of the thalamic sub-region in at least 5 of the 18 of the segmentation results from the normal subjects. Numbers in the top left corner of each image correspond to the slice number from Fig. 10.

Table 12
Fuzzy overlap OBL values for regions from the 18 normal datasets assigned to each of the six target regions. For each subject, the region sizes were adjusted for differences in the total intracranial volume of subjects before the mean for all subjects was calculated.

Region	Left thalamus		Region	Right thalamus	
	Fuzzy overlap (OBL)			Fuzzy overlap (OBL)	
	$\alpha_i=1$	$\alpha_i=1/V_i$		$\alpha_i=1$	$\alpha_i=1/V_i$
Frontal	0.508	0.500	Frontal	0.570	0.571
Occipital	0.168	0.128	Occipital	0.107	0.092
Parietal	0.469	0.459	Parietal	0.424	0.414
Post-central	0.205	0.189	Post-central	0.187	0.170
Pre-central	0.357	0.348	Pre-central	0.374	0.373
Temporal	0.494	0.494	Temporal	0.494	0.493

choice of the number of cortical targets was somewhat arbitrary, and was driven by a priori assumptions about cortical subdivisions. An alternative approach would be to use a data-driven technique to define thalamic regions based on connectivity profiles only, estimating the number of thalamic regions from the re-ordered cross-correlation matrix (Devlin et al., 2006; Johansen-Berg et al., 2004). Although this might be a more rigorous way to determine most likely patterns of thalamic subdivisions, here we are interested in exploring reproducibility of the data; hence we chose a direct and tractable approach using predetermined cortical segmentations to drive thalamic segmentation. Although Behrens et al. have proposed that these thalamic subdivisions correspond directly with thalamic nuclei; we prefer to make no such assumption, since there may be anatomically shared thalamo-cortical or cortico-thalamic pathways which emerge from, or project to, more than one adjacent thalamic nucleus. Johansen-Berg et al. (2005) showed that the method was reproducible between subjects when using the seven cortical regions defined in Behrens et al. (2003a). We have considerably extended their findings, showing the limits of expected reliability in a single subject, and demonstrating the potential to build probability maps based on larger subject numbers.

An alternative method to segment thalamus based on correlations between resting BOLD signal in thalamus and cortical regions has been proposed (Zhang et al., 2008). A recent comparison between this BOLD fMRI method and the method described here based on DTI has shown some correspondence between thalamic segmentations based on structural connections and based on functional connections (Zhang et al., 2010). However, this correspondence was not complete, suggesting each method may reveal different information about thalamic connectivity. A further alternative method to segment thalamus used spatial clustering of thalamic voxels according to T1 and T2 relaxation times in DESPOT imaging (Deoni et al., 2005, 2007). Here, we acquired DESPOT data in our subjects in the expectation that we would compare DTI-based and DESPOT-based segmentation methods. However, Deoni's method requires prior anatomical information regarding the relative size and approximate position of thalamic segments. At the time of writing, our exploration of this method suggests that apparent reliability is driven substantially by this prior anatomical information. Further investigation and optimisation of Deoni's method will be required before a reasonable comparison with DTI-based segmentation can be made.

Table 13
Summary of TAO values for analyses of single subject and multiple subject data.

	Number of target regions	Left thalamus		Right thalamus	
		$\alpha_i=1$	$\alpha_i=1/V_i$	$\alpha_i=1$	$\alpha_i=1/V_i$
Single dataset processed multiple times	6	0.949	0.934	0.966	0.908
	31	0.890	0.675	0.921	0.721
16 datasets of single subject	6	0.657	0.605	0.713	0.540
	31	0.462	0.215	0.520	0.189
18 normal subjects	6	0.407	0.287	0.430	0.280

Table 14
Summary of fuzzy overlap TAO values for analyses of single subject and multiple subject data.

	Number of target regions	Left thalamus		Right thalamus	
		$\alpha_i=1$	$\alpha_i=1/V_i$	$\alpha_i=1$	$\alpha_i=1/V_i$
Single dataset processed multiple times	6	0.941	0.924	0.960	0.919
	31	0.891	0.554	0.913	0.532
16 datasets of single subject	6	0.667	0.614	0.717	0.576
	31	0.501	0.236	0.551	0.204
18 normal subjects	6	0.451	0.336	0.469	0.330

Our segmentation results obtained using six large cortical target regions were similar to those obtained in previous studies. Initially, we examined one dataset, analysed multiple times, using a different seed for the random number generator each time. All segmentations of the thalami were very similar, as shown by the high OBP, OBL and TAO values. This result was expected, as the process should not be significantly influenced by the sequence of random numbers used. Increasing the number of target regions to 31 produced a similar pattern of connectivity to that obtained using six regions. However, further subdivision of the thalami was seen, particularly in the medial and posterior regions. Using 31 cortical target regions, there were some regions that had no strong connections to the thalamus (regions 11, 12, 23, 24 and 26–30 in both left and right thalami, regions 8, 19 and 25 in left thalamus and regions 5, 7, 18 in right thalamus). The TAO values found from the results using 31 regions were still very high, albeit slightly lower than those seen using six target regions. Using both six and 31 target regions, most voxels in the right and left thalami were classified as connecting to the same region in all 16 repeats of data processing.

When 16 datasets from the same subject were analysed using six target regions, similar patterns of connectivity were seen in all results. Compared to the repeated processing of the same dataset, more variability between the segmentation results was apparent, shown by reduced OBP and TAO values. As expected, most variability in the classification of the thalamic voxels occurred at the boundaries between the identified regions. The classification of thalamic voxels was more consistent in the lateral regions of the thalami, and more variability was observed in medial and posterior regions. Best overlap was observed for the target regions to which a larger region of the thalamus was classified as connecting (frontal, parietal, temporal and pre-central target regions). For the smaller thalamic regions, with connections to post-central and occipital target regions, overlap between the segmentation results from the 16 datasets was somewhat less strong, as shown by the lower OBL values particularly in the right thalamus. For these 16 datasets, increasing the number of target regions to 31 reduced the overlap, suggesting that if a comparison between subjects or subject groups were to be carried out, 31 target regions is likely to introduce too much variability. This might be relevant if group differences were sought in the medial or posterior regions.

For the 18 control datasets, a large degree of overlap in segmentation results was observed in regions classified as connecting to the frontal, parietal, temporal and pre-central target regions. For the other two target regions (post-central and occipital) only a few voxels were classified by the same label in more than half of the 18 results. These two regions were the smallest regions identified, so small differences in the location of these regions in the different subjects would have had a considerable effect on the overlap of the regions between subjects, as shown by the very low OBL values in both thalami. Results showed high similarity of the segmentation results for different subjects, particularly in the central portion of the regions identified. Reduced TAO values for the normal subjects, compared to results from multiple datasets from a single subject

suggested that there may be additional variation in the size or shape of the regions identified in different subjects. Some of this additional variation may be explained by the known inter-subject variability in the size and position of thalamic nuclei (Morel et al., 1997). Differences in thalamic segmentation might also have arisen due to slight differences in the cortical segmentation, particularly if paths from the thalamus progress toward a boundary region between target regions.

For individual subjects, this method of segmentation appeared to be most reproducible in specific regions of the thalamus, and for target regions to which a large region of the thalamus was classified as being connected. When a smaller thalamic region was classified as connecting to a particular target region, the resulting thalamic region was much less reproducible. It would be interesting to repeat this analysis using the multifibre approach described by Behrens et al. (2007) to determine whether improvements in reproducibility would be made, particularly for regions classified as connecting to post-central and occipital target regions. This analysis would however be most appropriate for diffusion imaging data acquired using a greater number of diffusion encoding directions than implemented in our data. It would be possible to extend our analyses in further ways: for example, as well as exploring the impact of altering the seeds for the random number generator in probtrackx, it would be possible similarly to alter the seed used in bedpostx. We do not expect this would lead to different conclusions from those reached here. It would also be valuable to explore between-subjects reproducibility using an additional measure which could take account of anatomical variability between subjects, such as Hausdorff distance.

In our examination of the overlap of hard-segmented labels, regions are assumed to be invariant in size and position between acquisitions in the same subject, or between subjects. It seems entirely reasonable to assume thalamic nuclear anatomy to remain unchanged between acquisitions carried out on the same healthy subject over a relatively short timeframe. However, it is not reasonable to assume thalamic nuclear anatomy to be identical between subjects, even in volumetrically normalised data. To address this issue, we incorporated an assessment of fuzzy overlap, which allows comparison of the connectivity distribution across the whole thalamus to each cortical target to be compared (between acquisitions in the same subject or between subjects). Comparison of Tables 5 and 6, and Tables 9 and 10, shows that in the single dataset analysed multiple times and in the multiple acquisitions in the single subject for 31 cortical targets, hard-segmented thalamic regions connecting to some targets had very little or no overlap, but the fuzzy overlap of connectivity distribution across the whole thalamus to these same targets was high. Since these were data from the same subject, we believe that this exposes unreliability of the segmentation method in certain thalamic regions, which in all instances correspond to smallest thalamic hard segments. In contrast, comparison of Tables 3 and 4, and Tables 7 and 8 shows that in the single dataset analysed multiple times and in the multiple acquisitions in the single subject for 6 cortical targets, there was only a marginal improvement in fuzzy overlap measures compared with overlap of hard segments (of the order of 5% on average). However, comparison of Tables 11 and 12, which show overlap measures for 18 subjects, reveals that fuzzy overlap measures are on average 23% higher than “hard segment” overlap measures, suggesting that there is a degree of variability in anatomical position and size of thalamic nuclei between subjects. We conclude that the method described here reliably segments the thalamus in multiple acquisitions in the same subject; and reliably locates the size and position of thalamic segments between subjects; and is able to capture between-subjects variability that we propose stems from real anatomical differences between subjects.

As methods such as this become translated into the clinical domain, it will be important to be aware that reliability and reproducibility of target identification will vary between nuclei, and that this might need to be subjected to careful evaluation for each nucleus and each clinical implementation. Additionally, it will of course be essential to undertake reliability and reproducibility studies of this kind in the relevant patient groups. We anticipate that future studies will concentrate on developing a probabilistic atlas of thalamic connectivity based on a large number of normal subjects of varying age; and on developing similar atlases for relevant diseases. We also suggest that future work may cast light on whether regions identified using connectivity-based thalamic segmentation correspond directly to thalamic nuclei. Our finding of good intra-subject reproducibility, coupled with a larger inter-subject variability, suggests that thalamic division by probabilistic tractography represents a reliable measurement. The finding of inter-subject variability should therefore allow meaningful correlations of measures of thalamic connectivity or volumes of thalamic subdivisions with clinical variables.

Acknowledgments

MPR is supported by the Getty Family Foundation. WRC acknowledges financial support from the Department of Health via the National Institute for Health Research (NIHR) specialist Biomedical Research Centre in Mental Health award to South London and Maudsley NHS Foundation Trust. RAH was supported by a grant from the Dunhill Medical Trust.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2010.04.024.

References

Ahsan, R.L., Allom, R., Gousias, I.S., Habib, H., Turkheimer, F.E., Free, S., Lemieux, L., Myers, R., Duncan, J.S., Brooks, D.J., Koepp, M.J., Hammers, A., 2007. Volumes, spatial extents and a probabilistic atlas of the human basal ganglia and thalamus. NeuroImage 38, 261–270.

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

Behrens, T.E., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C.A., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M., 2003a. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat. Neurosci. 6, 750–757.

Behrens, T.E., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., Matthews, P.M., Brady, J.M., Smith, S.M., 2003b. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magn. Reson. Med. 50, 1077–1088.

Behrens, T.E.J., Berg, H.J., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W., 2007. Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? NeuroImage 34, 144–155.

Ciccarelli, O., Catani, M., Johansen-Berg, H., Clark, C., Thompson, A., 2008. Diffusion-based tractography in neurological disorders: concepts, applications, and future developments. Lancet Neurol. 7, 715–727.

Crum, W.R., Camara, O., Hill, D.L., 2006. Generalized overlap measures for evaluation and validation in medical image analysis. IEEE Trans. Med. Imaging 25, 1451–1461.

Deoni, S.C., 2007. High-resolution T1 mapping of the brain at 3T with driven equilibrium single pulse observation of T1 with high-speed incorporation of RF field inhomogeneities (DESPOT1-HIFI). J. Magn. Reson. Imaging 26, 1106–1111.

Deoni, S.C., Josseau, M.J., Rutt, B.K., Peters, T.M., 2005. Visualization of thalamic nuclei on high resolution, multi-averaged T1 and T2 maps acquired at 1.5 T. Hum. Brain Mapp. 25, 353–359.

Deoni, S.C., Rutt, B.K., Parrent, A.G., Peters, T.M., 2007. Segmentation of thalamic nuclei using a modified k-means clustering algorithm and high-resolution quantitative magnetic resonance imaging at 1.5 T. NeuroImage 34, 117–126.

Devlin, J.T., Sillery, E.L., Hall, D.A., Hobden, P., Behrens, T.E.J., Nunes, R.G., Clare, S., Matthews, P.M., Moore, D.R., Johansen-Berg, H., 2006. Reliable identification of the auditory thalamus using multi-modal structural analyses. NeuroImage 30, 1112–1120.

Gousias, I.S., Rueckert, D., Heckemann, R.A., Dyet, L.E., Boardman, J.P., Edwards, A.D., Hammers, A., 2008. Automatic segmentation of brain MRIs of 2-year-olds into 83 regions of interest. NeuroImage 40, 672–684.

Hammers, A., Allom, R., Koepp, M.J., Free, S.L., Myers, R., Lemieux, L., Mitchell, T.N., Brooks, D.J., Duncan, J.S., 2003. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. Hum. Brain Mapp. 19, 224–247.

Heckemann, R.A., Hajnal, J.V., Aljabar, P., Rueckert, D., Hammers, A., 2006. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. NeuroImage 33, 115–126.

Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. Med. Image Anal. 5, 143–156.

Johansen-Berg, H., Behrens, T.E., Robson, M.D., Drobnjak, I., Rushworth, M.F., Brady, J.M., Smith, S.M., Higham, D.J., Matthews, P.M., 2004. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. Proc. Natl. Acad. Sci. U. S. A. 101, 13335–13340.

Johansen-Berg, H., Behrens, T.E., Sillery, E., Ciccarelli, O., Thompson, A.J., Smith, S.M., Matthews, P.M., 2005. Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. Cereb. Cortex 15, 31–39.

Jones, D.K., Griffin, L.D., Alexander, D.C., Catani, M., Horsfield, M.A., Howard, R., Williams, S.C., 2002a. Spatial normalization and averaging of diffusion tensor MRI data sets. NeuroImage 17, 592–617.

Jones, D.K., Williams, S.C.R., Gasston, D., Horsfield, M.A., Simmons, A., Howard, R., 2002b. Isotropic resolution diffusion tensor imaging with whole brain acquisition in a clinically acceptable time. Hum. Brain Mapp. 15, 216–230.

Koller, W., Pahwa, R., Busenbark, K., Hubble, J., Wilkinson, S., Lang, A., Tuite, P., Sime, E., Lazano, A., Hauser, R., Malapira, T., Smith, D., Tarsy, D., Miyawaki, E., Norregaard, T., Kormos, T., Olanow, C.W., 1997. High-frequency unilateral thalamic stimulation in the treatment of essential and parkinsonian tremor. Ann. Neurol. 42, 292–299.

Morel, A., Magnin, M., Jeanmonod, D., 1997. Multiarchitectonic and stereotactic atlas of the human thalamus. J. Comp. Neurol. 387, 588–630.

Nowinski, W.L., Liu, J., Thirunavuukarasuu, A., 2008. Quantification and visualization of three-dimensional inconsistency of the ventrointermediate nucleus of the thalamus in the Schaltenbrand–Wahren brain atlas. Acta Neurochir. (Wien) 150, 647–653 discussion 653.

Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans. Med. Imaging 18, 712–721.

Schaltenbrand, G., Wahren, W., 1977. Atlas for stereotaxy of the human brain. Thieme, Stuttgart.

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

Theodore, W.H., Fisher, R.S., 2004. Brain stimulation for epilepsy. Lancet Neurol. 3, 111–118.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation–maximization algorithm. IEEE Trans. Med. Imaging 20, 45–57.

Zhang, D., Snyder, A.Z., Fox, M.D., Sansbury, M.W., Shimony, J.S., Raichle, M.E., 2008. Intrinsic functional relations between human cerebral cortex and thalamus. J. Neurophysiol. 100, 1740–1748.

Zhang, D., Snyder, A.Z., Shimony, J.S., Fox, M.D., Raichle, M.E., 2010. Noninvasive Functional and Structural Connectivity Mapping of the Human Thalamocortical System. Cereb Cortex 20, 1187–1194.