Medical Image Analysis 12 (2008) 713–726

Contents lists available at ScienceDirect

Medical Image Analysis

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

Towards a validation of atlas warping techniques

M. Mallar Chakravarty [a,*], Abbas F. Sadikot [a,b], Jürgen Germann [a], Gilles Bertrand [a,b], D. Louis Collins [a]

[a] McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University 3801, University Street, Montréal, Canada H3A 2B4
[b] Division of Neurosurgery, McGill University, Montréal, Canada

ARTICLE INFO

Article history:
Received 29 May 2007
Received in revised form 23 April 2008
Accepted 28 April 2008
Available online 13 May 2008

Keywords:
Movement disorders
Parkinson's disease
Validation
Nonlinear registration
Subcortical atlas
Surgical atlases
Digital atlases
Atlas warping

ABSTRACT

Pre-operative magnetic resonance (MRI) and computed tomography (CT) image volumes are often used for planning and guidance during functional neurosurgical procedures. These operations can include the creation of lesions in the thalamus (thalamotomy) or the globus pallidus (pallidotomy), or the insertion of deep brain stimulation (DBS) electrodes in the subcortical nuclei. These subcortical targets are often difficult to localize in pre-operative imaging data due to the limited resolution and contrast available in standard MRI or CT techniques. To address this problem, digital atlases of subcortical nuclei are often used to accurately identify surgical targets since they can be warped to fit each patient's unique anatomy. Targeting accuracy thus depends on the quality of the atlas-to-patient warp.

In this paper, three atlas-to-patient warping techniques are compared. Two methods rely on an MRI template as an intermediary to estimate a nonlinear atlas-to-patient transformation. The third is novel, and uses a pseudo-MRI derived from an atlas of the basal ganglia and thalamus to estimate the nonlinear atlas-to-patient transformation directly. The methods are compared using (1) manual segmentations of subcortical nuclei and (2) functional data from intra-operative thalamic stimulation. The results demonstrate that the template-based atlas-to-patient warping technique is the best of the three for customizing the atlas onto patient data.

1. Introduction

Movement disorders are often treated with functional neurosurgical procedures. These procedures can include the creation of lesions in the thalamus (thalamotomy) (Atkinson et al., 2002; Duval et al., 2005; Lenz et al., 1995; Otsuki et al., 1994) the globus pallidus (pallidotomy) (Cohn et al., 1998; Gross et al., 1999; Lombardi et al., 2000; Starr et al., 1999) as well the introduction deep brain stimulation (DBS) electrodes in the thalamus, globus pallidus, or the subthalamic nucleus (STN) (Eskandar et al., 2001; Krause et al., 2001; Sanchez Castro et al., 2006b; Starr et al., 1999). These procedures require accurate target localization of these subcortical structures using pre-operative magnetic resonance (MRI) or computed tomography (CT) image volumes. Despite recent advances in medical imaging techniques which allow improved visualization of the thalamus (Behrens et al., 2003; Deoni et al., 2005; Johansen-Berg et al., 2005), most clinical MRI volumes lack the contrast and resolution required to properly visualize the subcortical nuclei. These pre-operative scans are often taken with the body coil since a stereotactic head-frame is attached to the patient's head to establish a co-ordinate system within the image volume. In addition, a variety of other factors can limit contrast and signal-to-noise ratio in awake patients undergoing stereotactic surgery, including motion artefacts related to movement disorders, and the limits of MRI acquisition time. These limitations can cause difficulty for surgeons when distinguishing between the borders of different subcortical nuclei used as surgical targets (see Fig. 1). Originally, print atlases were used in conjunction with anatomical landmarks to guide functional neurosurgical procedures (Schaltenbrand and Wahren, 1977; Talairach and Tournoux, 1988). However, digital atlases have proven to be useful in surgical planning and guidance as they can be customized to pre-operative patient data (Bertrand et al., 1973; Chakravarty et al., 2006a; Finnis et al., 2003; Ganser et al., 2004; Nowinski et al., 2000; St-Jean et al., 1998).

1.1. Computerized atlases for surgical planning

The use of *computerized atlases* to guide functional neurosurgery was pioneered at the Montreal Neurological Institute in work done by Bertrand (1982), and Bertrand et al. (1973, 1974). This work used a digitized version of the Schaltenbrand and Bailey atlas (Schaltenbrand and Bailey, 1959) and used an intra-operative ventriculogram as an imaging reference. Hardy, Bertrand and Thompson later used this atlas in a comprehensive analysis of the somatopy of fiber tracts and subcortical nuclei in the human brain in the late 1970s and early 1980s (Hardy et al., 1979a,b,c,d,e, 1980a,b,c, 1981).

* Corresponding author. Tel.: +1 514 398 1730; fax: +1 514 398 2975.
E-mail address: mallar@bic.mni.mcgill.ca (M.M. Chakravarty).

1361-8415/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.media.2008.04.003

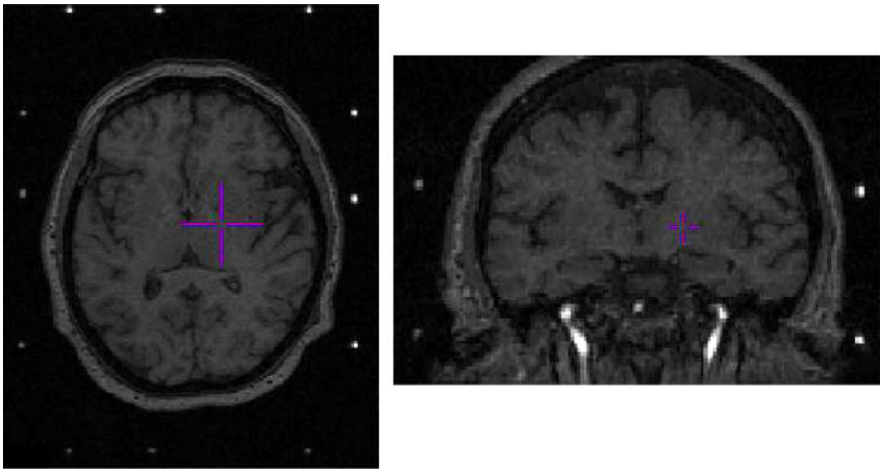

Fig. 1. Example of pre-operative MRI volume with head-frame affixed to patient. (Left) Sagittal view; (middle) coronal view; (right) axial view. Image volume shows lack of contrast in the sub-cortical nuclei. Surgical targeting is extremely difficult in these nuclei as a result.

More recently, several atlases of the human brain have been used to guide stereotaxic procedures. Several of these atlases rely on digitized versions of the print atlases which are commonly available. Nowinski et al. (1997) have developed a digital atlas that incorporated data from three print atlases including Ono et al. (1990), Schaltenbrand and Wahren (1977), and Talairach and Tournoux (1988). In order to register the atlas to a subject or patient, a piece-wise linear approach is used to transform the atlas to the MR volume. Later, the atlas-to-patient warping approach was refined by estimating a nonlinear transformation using a finite element approach (Xu and Nowinski, 2001). The Talairach and Tournoux atlas was also used as the foundation for a digital atlas by Ganser et al. (2004). The original plates were scanned and reconstructed in three dimensions by estimating a Delaunay tetrahedrization on each structure of the atlas. Surfaces of these structures were extracted using the marching cubes algorithm (Lorensen and Cline, 1987). Their atlas warping technique improves on Talairach and Tournoux's (1988) so-called "proportional grid system" by using a set of radial basis functions to define a deformation field.

Many groups use template based procedures to warp atlases to pre-operative patient data. These procedures warp an atlas to a template MRI. After *atlas-to-template* alignment, the atlas warping problem is narrowed down to a standard MRI-to-MRI registration problem. Once the template-to-patient transformation is estimated it is applied to the atlas to customize it to patient data. For these template-based techniques, the atlas-to-patient and template-to-patient warps are equivalent.

In the original work of St-Jean et al. (1998) the Colin27 MRI template (Holmes et al., 1998) was used to estimate atlas-to-patient transformations. This template is the average of 27 T1-weighted MRI volumes of the same subject which results in a volume with enhanced signal-to-noise ratio and improved contrast. A 3D version of the Schaltenbrand and Wahren atlas was warped to the template using a thin-plate-spline (TPS) warp (Bookstein, 1989). The TPS warp is a landmark-based technique which aligns landmarks exactly and will interpolate the data between these landmarks by constraining them to the properties of a thin metal plate. As a result, it is possible that there are some misregistrations in the initial atlas-to-template customization, and these errors will be propagated throughout the atlas-to-patient warping process. This is the atlas previously used at the Montréal Neurological Institute.

The Colin27 high-resolution MRI template (Holmes et al., 1998) has also been used to estimate nonlinear atlas-to-patient transformations by Finnis et al. (2003). In their work, a *probabilistic functional atlas* was developed by warping a digitized version of the Schaltenbrand and Wahren atlas (Schaltenbrand and Wahren, 1977). Co-ordinates representing successful intra-operative stimulation points were clustered into the common atlas space using nonlinear registration methods. To customize the atlas to patient data, the nonlinear transformation was then applied to the digitized atlas and the point clusters. Recent work has expanded the amount of intra-operative electro-physiological data that has been included in their electro-physiological database (Guo et al., 2005).

D'Haese et al. have developed an atlas (D'Haese et al., 2005a) which uses electro-physiological intra-operative recordings from the STN registered to a template created from the average of pre-operative data. This choice of template was later refined (D'Haese et al., 2005b) to a single MRI template, selected based on the best prediction of the final DBS location when using radial basis functions to estimate the nonlinear template-to-patient transformation (Rohde et al., 2003).

Bardinet et al. (2005) reconstructed a set of serial histological data and warped this data set to T1 and T2 reference MRIs using only linear transformations. The transformations are optimized at the level of the basal ganglia and thalamus. Contours of structures in the basal ganglia and the mesencephalon were manually traced on the histology to allow for increased visualization of the anatomy. The final atlas-to-patient transformation was estimated using an affine transformation estimated in a hierarchical fashion and is used to predict the location for the STN DBS stimulator. This position was then correlated with intra-operative electro-physiological findings. The methods used for the development of this atlas were recently detailed in Yelnik et al. (2007).

The use of expert identification of surgical targets was used to create an atlas for STN DBS targeting by Sanchez Castro et al. (2005, 2006a,b). This atlas is based on the manual identification of the DBS target on multiple patients. The target labeling is repeated on all subjects, by the same expert on five separate occasions to eliminate intra-rater error. The points from each subject were averaged and then warped to a reference pre-operative MRI in order to create an atlas. Four different warping techniques were evaluated (Schaltenbrand and Wahren (1977), affine transformations (Maes et al., 1997), Demons algorithm (Thirion et al., 1996; Thirion, 1998), and B-splines (Rueckert et al., 1999)) using correspondence to the final target point assessed in post-operative data as the evaluation criteria.

1.2. Atlas warping validation

Each procedure described above is predicated on finding and optimizing techniques for atlas-to-patient warping. The goal is to integrate these atlases in the planning phase for functional neurosurgery. Thus the pre-operative planning is limited by the accuracy of the warping techniques used for atlas-to-template warping. However validation of any warping techniques is typically difficult due to the lack of an universally accepted "gold standard" and requires the development of novel techniques for evaluation. Many of the groups have used intra-operative functional recordings (Bardinet et al., 2005), post-operative data (D'Haese et al., 2005a), or STN DBS targets defined by surgical experts (Sanchez Castro et al., 2006a, 2005). To our knowledge, groups have not used a volumetric representation of the anatomy for validation. Finnis et al. (2003) used intra-operative stimulation eliciting visual activations and compared the location of these stimulation points with regards to their position relative to the optic tract. As an extension to this original work, Guo et al. (2006) recently published a more comprehensive validation of the initial work from this group using final target locations from pallidotomies, thalamotomies, and thalamic and subthalamic nucleus DBS implantations. In the validation provided by Nowinski et al. (2000) for the warping techniques for their multiple brain atlas database, a comparison was done of subjects who were operated (for pallidotomy, thalamotomy, and thalamic DBS) and the final target location was compared with respect to the atlas-derived predicted target location. However, only visual validation results were presented when the atlas-to-patient and inter-atlas (i.e., Talairach to Schaltenbrand atlas) warping techniques was refined using an FEM method (Xu and Nowinski, 2001). In the atlas developed by Ganser et al. (2004) only the location of the frontal tip of the putamen was verified for validation purposes. However the accuracy of the identification of structures is critical as targets are often chosen relative to structures which are easily visible using pre-operative MRI volumes.

In other comparisons of nonlinear registration algorithms, the use of volumetric classifications and segmentations have been used optimize or compare different strategies. Robbins et al. (2004) used the minimization of entropy between segmented MRI volumes as a technique for the optimization of nonlinear registration parameters of the automatic nonlinear image matching and anatomical labeling (ANIMAL) algorithm of Collins and Evans (1997), and Collins et al. (1995). In a broader study Hellier et al. (2003) studied the accuracy of several commonly-used nonlinear registration techniques using several different criteria including: global volume, overlap of different segmented tissue classes, curvature of the iso-intensity surfaces, consistency of the nonlinear deformation, as well as quantitative and qualitative evaluation of sulcii after nonlinear warping.

1.3. Goals

In this paper we use an atlas created from reconstructed serial histological data (Chakravarty et al., 2006a) which has been warped to the Colin27 high resolution MRI template (Holmes et al., 1998) (based on an *atlas-to-template* warping procedure). This anatomical atlas is described in Section 2. The atlas is customized to a particular patient's anatomy (using an *atlas-to-patient* warping procedure) with the ANIMAL nonlinear registration technique (Collins and Evans, 1997; Collins et al., 1995) (Section 2.2), using three different strategies: two depend on the Colin27 template and the third uses a "pseudo-MRI" (Chakravarty et al., 2006a). The work presented here is a continuation of our preliminary validation work (Chakravarty et al., 2005).

The main goal of this paper is to compare three atlas-to-patient warping strategies to determine which is best for surgical planning. To achieve this goal, three subgoals must be met: (1) validate the initial atlas-to-template warping procedure, since any errors in atlas–template alignment will be propagated in the two template-based atlas-to-patient mappings; (2) develop metrics for minimally-biased anatomical evaluation of the atlas-to-patient warping techniques; (3) develop metrics for electro-physiological evaluation of the atlas-to-patient warping techniques. In addition to validating the ANIMAL algorithm for atlas customization in the context of targeting in functional neurosurgical procedures used in movement disorders, the contributions of this paper include the development and evaluation of a pseudo-MRI-based nonlinear registration procedure, the development of manually-derived silver-standard labels for anatomical validation, and the development of new metrics based on electro-physiological recordings for further warping validation.

The paper is organized in the following manner: Section 2 presents the methods, describing both the atlas (Section 2.1) and the nonlinear registration algorithm (Section 2.2) used. Section 2.3 describes the atlas-to-template warping strategy, while Sections 2.4–2.6 describe the three atlas-to-patient warping strategies. A series of experiments are described in Section 3. The creation and evaluation of the manual label-based silver standard is described in Section 3.2. Once the quality of the silver standard has been characterized (Section 3.2.3), it is used to evaluate the atlas-to-template warping (Section 3.2.4) and the three atlas-to-patient warping strategies (Section 3.2.5). Section 3.3 describes the electro-physiological metrics and validation. Section 4 presents the experimental results, Section 5 includes a summary and discussion and Section 6 ends with conclusions and suggestions for future work.

2. Methods

In this section the atlas of the basal ganglia and thalamus developed for planning functional neurosurgical procedures and the techniques used to warp this atlas to a high resolution MRI template are discussed in Section 2.1. The ANIMAL algorithm used for atlas-to-template and atlas-to-patient nonlinear warping is described in Section 2.2. The atlas-to-template warping procedure is described in Section 2.3. The three different atlas-to-patient warping techniques based on the ANIMAL algorithm are described in Sections 2.4–2.6.

In order to properly identify subcortical structural anatomy, the atlas must be warped to effectively match the patient's subcortical morphology. The three warping techniques described here start with the estimation of a linear nine parameter (three translations, three rotation, and three scaling factors) (Collins et al., 1994) transformation used to map the atlas into the native space of the MRI data. This transformation is used as the input for the nonlinear transformation estimation, where it is refined with the ANIMAL algorithm. The ANIMAL algorithm is used to estimate the two different template based warping procedures and the single, direct, pseudo-MRI based techniques (all of which are evaluated in this paper). Standard execution times for a Apple MacBook G4 2 GHz Intel Core 2 Duo, with 2 GB of 667 MHz DDR2 SDRAM are given in Sections 2.4–2.6.

2.1. Computerized atlas of the basal ganglia and thalamus

In this paper we use a bi-lateral version of a new high resolution atlas that contains multiple registered representations of 105 subcortical grey and white matter structures (Chakravarty et al., 2006a). The data is derived from 84 sections of manually segmented serial histological data. The atlas contains nomenclature from three different sources for gross-anatomy (Schaltenbrand and Wahren, 1977), the thalamus (Hirai and Jones, 1989), and the temporal lobe

(Gloor, 1997). The histological data was reconstructed using optimized nonlinear slice-to-slice morphological and intensity corrections techniques (Chakravarty et al., 2006a). This atlas has also been used to create a probabilistic functional atlas using post-operative active electrode data (Chakravarty et al., 2006b).

A voxel-label-atlas was created by assigning each structure a unique label which acts an anatomical identifier. To estimate the nonlinear atlas-to-template transformation, a pseudo-MRI (to facilitate the atlas-to-template warping procedure described in Section 2.3) was created by modifying the intensities of the voxel-label-atlas to match those of the Colin27 MRI template (Holmes et al., 1998).

The voxel-label atlas can also be viewed as a set of 3D objects, thus aiding to understand the 3D relationships between different structures and can be navigated in register with the atlas (Chakravarty et al., 2006a). The four different atlas representations (histological, voxel-label, pseudo-MRI) can be viewed together depending on user preference. These representations, along with the Colin27 MRI template can be seen in Figs. 2 and 3.

2.2. The ANIMAL algorithm

The ANIMAL algorithm is an iterative procedure which estimates a 3D deformation field that matches a *source volume* to

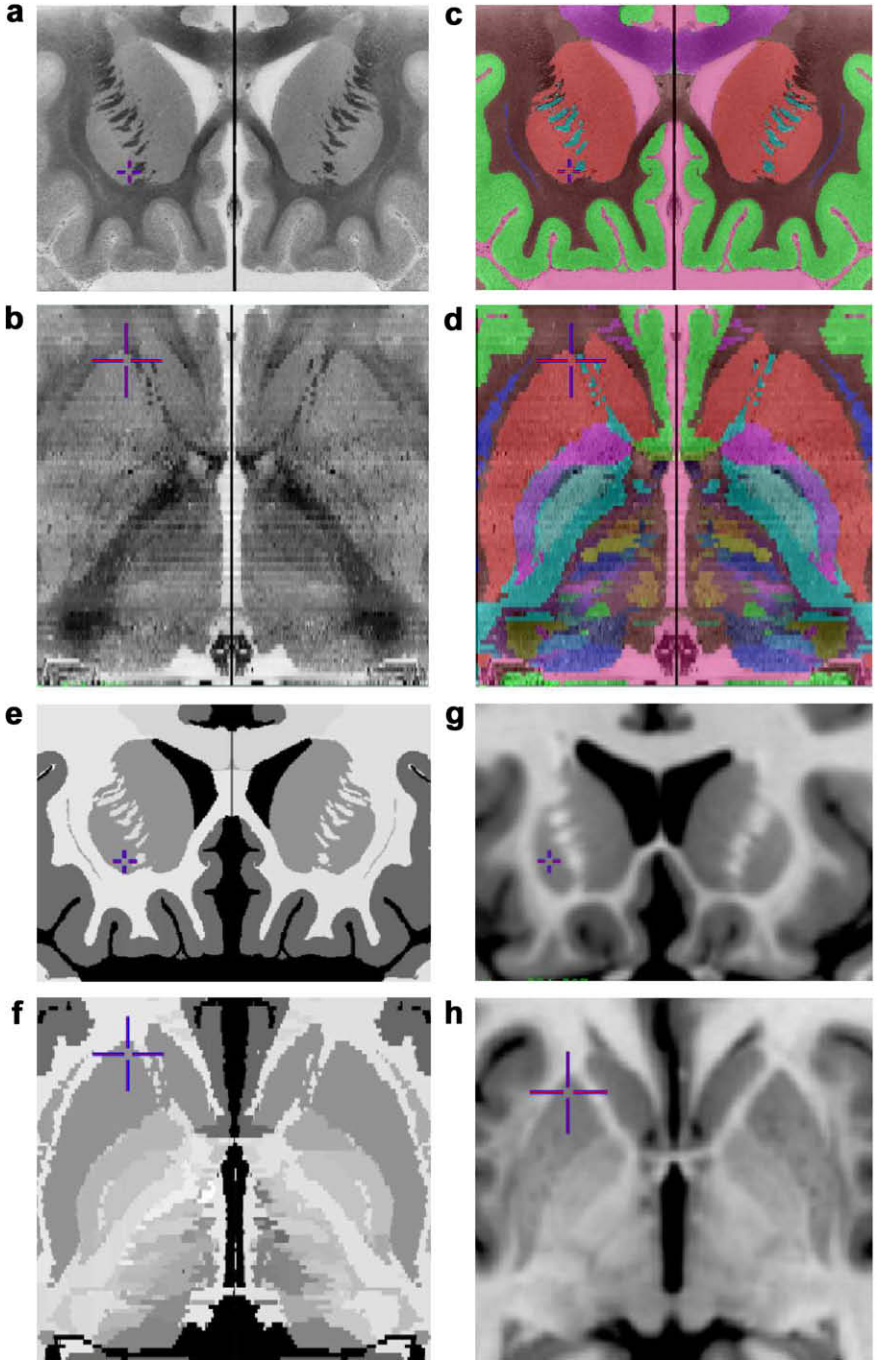

Fig. 2. Coronal and axial representation of atlas described in Section 2.1. (a) Original coronal section from histological dataset. (b) Reconstructed transverse slice through histological volume. (c) and (d) Voxel-label-atlas representation of the atlas. (e) and (f) Pseudo-MRI representation of the atlas. (g) and (h) Close-up of the Colin27 MRI template in the region of the basal ganglia and thalamus.

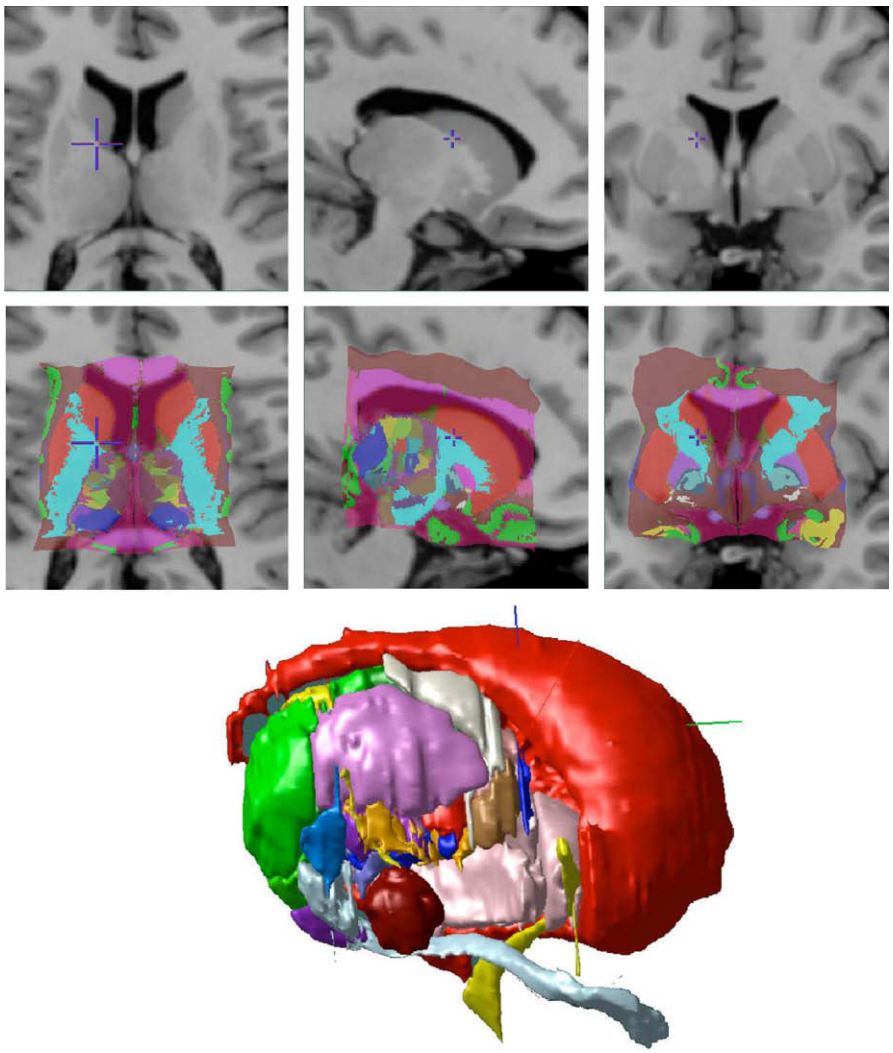

Fig. 3. (From top to bottom) Axial, sagittal, and coronal views of the Colin27 MRI template and the atlas warped to the Colin27 template. The 3D geometric atlas is shown on the bottom of the figure. The top two rows of this figure also demonstrate the extents of the ROIs used for atlas warping Methods B and C.

a *target volume*. The algorithm is divided into two steps. The first is the *outer loop*, where large deformations are estimated on data which has been blurred using a Gaussian kernel with a large full-width-at-half-maximum (FWHM). These larger deformations are then input to subsequent steps where the fit is refined by estimating smaller deformations on data blurred with a smaller FWHM.

At each step of the outer loop, the ANIMAL algorithm is applied iteratively in an *inner loop* to optimize the nonlinear transformation (N) that maximizes the similarity between a source volume (S) and a target volume (T) with the following objective function $\mathbf{\Gamma}$:

$$\mathbf{\Gamma}(S,T;N) = \boldsymbol{\beta}(S,T;N) + \mathbf{C}(N), \tag{1}$$

where $\boldsymbol{\beta}$ is the local similarity measure (i.e., the correlation ratio) and $\mathbf{C}$ is the cost function. The cost function yields large values for large deformations and smaller values for smaller deformations (effectively penalizing the objective function when large deformations are estimated).

The nonlinear transformation is represented by a deformation field that is iteratively estimated in a two step process: the first step involves the estimation of local translations for each node defined by optimizing equation (1) and the second is a smoothing step to ensure that the deformation field is continuous and does not cause stretching, tearing, or overlap. Three parameters can be set which help define the quality of the nonlinear transformation: the *similarity* (which balances the objective function with the cost function), the *weight* (which determines the proportion of each local translation estimated at one iteration that will be used at the next iteration), and the *stiffness* (which determines the smoothness of the nonlinear deformation field). For the warping techniques described Sections 2.3–2.6, the *similarity, weight*, and *stiffness* are all set to the same value (0.3, 1, and 1, respectively) according to the parameter optimization of Robbins et al. (2004). The reader is referred to Chakravarty et al. (2006a), Collins and Evans (1997), Robbins et al. (2004) for more details on these parameters.

At each step of the outer loop, the inner loop is run over a fixed number of iterations. Here the deformation field is defined by a set of local translations defined on a grid of equally spaced nodes. The resolution of the deformation field can be defined by setting three parameters: *step*, *lattice diameter*, and *sub-lattice*. The *step* parameter defines the space between each node in the lattice. The *lattice-diameter* is the diameter of the local spherical neighborhood around each node in which the objective function is estimated. The *sub-lattice* is the number of elements used within that local neighborhood. The values for these parameters are defined in Sections 2.4–2.6.

2.3. Atlas-to-template warping

Using the ANIMAL algorithm (Collins et al., 1995; Collins and Evans, 1997), we have improved on the atlas-to-template warping presented in the original digital atlas used at the Montréal

Neurological Institute developed by St-Jean et al. (1998). Here ANIMAL is used to estimate the nonlinear transformation required to map the pseudo-MRI described in Section 2.1 (source volume) to the Colin27 MRI template (target volume) (Chakravarty et al., 2006a). This is an improvement on the TPS warp estimated in the original work of St-Jean et al. (1998).

The transformation estimated here uses a hierarchical fitting strategy, however no blurring of either the pseudo-MRI or the Colin27 MRI template is done. Effective blurring is achieved due to the homogeneous nature of the pseudo-MRI intensities and by the sub-sampling done in estimation of the sub-lattice described in Section 2.2. A region of interest (ROI) is used to constrain the estimation of the deformation field to the sub-cortical region. The ROI is defined laterally by the insula, anteriorly by the genu of the corpus callosum, posteriorly by the limits of the later ventricles, superiorly by cingulate gyrus, and inferiorly by the temporal poles and can be seen in Fig. 3. The final atlas-to-template transformation is estimated on a dense 1 mm grid. The registration parameters are given in Table 1. By using this transformation to warp the voxel-label and 3D geometric atlases to the Colin27 MRI template we afford the user enhanced anatomical visualization (see Fig. 3).

2.4. Template based atlas-to-patient warping: Method A

Since the atlas and the Colin27 template are co-localized, the first atlas-to-patient warping technique is really a template-to-patient warping problem and is treated as a standard MRI-to-MRI registration problem. The registration parameters used were optimized and discussed in Robbins et al. (2004). This technique estimates the nonlinear transformation in a hierarchical fashion as explained in Section 2.2. The source volume is the Colin27 MRI template and the target is the patient MRI volume. The final transformation estimated is defined by vectors on a grid of equally spaced nodes that are 2 mm apart and applied to the original atlas warped to the Colin27 MRI template to complete the atlas-to-patient warp. The standard parameters used for the registration strategy are shown in Table 2. Execution time for this method was approximately 50 min.

2.5. Template based atlas-to-patient warping: Method B

Similar to Method A, the second template-based technique uses ANIMAL to estimate a template-to-patient nonlinear transformation. This technique is also hierarchical and the data is not blurred. Since Method B is inspired by the initial atlas-to-template warp, there are two main differences with Method A. While the Colin27 template is still used as the source volume, a cropped volume ROI volume (as described in Section 2.3) focused on the subcortical nuclei is used to limit the transformation estimation to facilitate the computational burden of estimating a higher-resolution transformation. The hierarchical nonlinear transformation estimation of this technique outputs a final transformation defined by vectors on a grid of equally spaced nodes that are 1 mm apart. The parameters for ANIMAL used for this template-based transformation estimation is shown in Table 1, and are the same one used for the atlas-to-template warp. We hypothesize that Method B will give better results than Method A due to the higher resolution deformation. Execution time for this method was approximately 70 min.

Table 1
ANIMAL parameters used for high-resolution template-based and pseudo-MRI based atlas-to-subject nonlinear transformation estimation

Step	Step size (mm)	Sub-lattice diameter	Sub-lattice	Iterations
1	4	8	6	15
2	2	6	6	15
3	1	6	3	15

Table 2
ANIMAL parameters used for template based atlas-to-subject nonlinear transformation estimation

FWHM (mm)	Step size (mm)	Sub-lattice diameter	Sub-lattice	Iterations
8	8	24	6	30
8	4	12	4	30
4	2	6	6	10

2.6. Pseudo-MRI based atlas-to-patient warping: Method C

Method C is based on the work in Chakravarty et al. (2006a, 2005) where the pseudo-MRI (described is Section 2.1) was used to directly estimate the atlas-to-template transformation used to warp the atlas directly to the Colin27 MRI template. Since the pseudo-MRI was created by modifying the intensities of the voxel-label-atlas to approximate the contrast of the Colin27 template, it can be used directly to estimate the atlas-to-patient transformation (rather than use the Colin27 as an intermediate template). The transformation estimation process uses the pseudo-MRI as the source volume and the patient MRI as the target volume, and estimates a nonlinear transformation using the ANIMAL algorithm. The parameters for ANIMAL are the same as those used in Method B, thus the final transformation estimated is defined by vectors on a grid with 1 mm spacing. The nonlinear transformation estimation procedure can be seen in the top path of the flowchart in Fig. 4 and the template based techniques are represented at the bottom of the same figure. The registration parameters are shown in Table 1 and are those that are used for the initial atlas-to-template transformation estimation. We hypothesize that Method C will yield better results than Method A or B because errors in atlas-to-template warping (Section 2.3) are eliminated as the transformation between atlas and patient are estimated directly. Execution time for this method was approximately 65 min.

3. Atlas warping evaluation – experimental methods

The following sections present anatomical and functional validation for the atlas-to-patient transform using manual anatomical labels and intra-operative electro-physiological recordings. The manual labeling protocol is first presented followed by four different validation experiments. In the first experiment, the quality of the manual labels are evaluated to verify that they can be used to validate both the atlas-to-template and atlas-to-patient warps. The second experiment evaluates the quality of the atlas-to-template match using the anatomicals labels. The third experiment also uses the manual labels to evaluate the atlas-to-patient customizations. Finally, the fourth experiment evaluates the atlas-to-patient customizations using the electro-physiological recordings.

3.1. Subjects

The atlas-to-patient warping techniques presented in Sections 2.4–2.6 were evaluated using clinical T1-weighted pre-operative image data from nine patients who had undergone thalamotomies (five males and four females, five left and four right thalamotomies). All MRIs were taken between 1997 and 2002 with the stereotactic head-frame attached on a Philips 1.5 T MRI scanner with axial slices with a 1 mm in plane voxel spacing and 1.5 mm voxel spacing between slices. Since the head-frame cannot fit inside the headcoil, scans were acquired in the body coil. Informed consent was obtained from all subjects involved in this study,

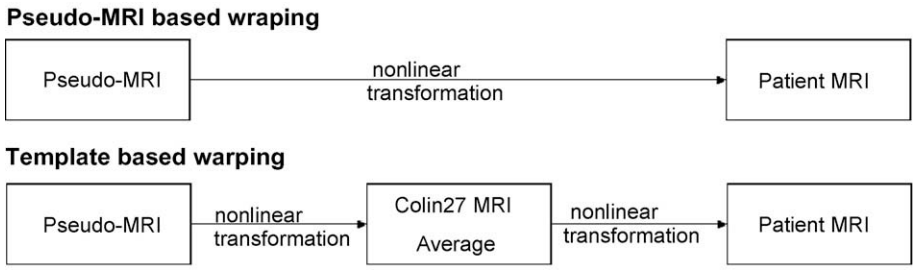

Fig. 4. Flow chart demonstrating the difference between Method C and Methods A and B technique. The top path, demonstrates how Method C bypasses the need to register the atlas to a patient MRI volume by directly estimating an atlas-to-patient transformation (see Section 2.6). The bottom path shows the template based techniques of Methods A and B (see Sections 2.4 and 2.5).

and the research protocol was approved by the ethics board of the Montreal Neurological Hospital and Institute.

3.2. Anatomical validation

Since a gold standard for anatomical validation was not available, manual structure segmentations were used to evaluate the "goodness of fit" of the atlas-to-template and atlas-to-patient warping techniques. To estimate inter-rater variability, five expert raters identified the striatum (the caudate nucleus, the putamen, and the nucleus accumbens), the thalamus, and the globus pallidus bilaterally in the patient MRI data and on the Colin27 template. The striatum was used as it is a large high contrast structure; this feature facilitate its identification through manual labeling. Furthermore, surgeons use the location of the putamen to help guide entry towards the STN, as it has been hypothesized the micro-lesions in this structure may be linked to post-mortem morbidity (Benabid et al., 2002). The thalamus and the globus pallidus are still targets for lesioning and stimulations for Parkinsons disease (PD). Beyond PD, the thalamus is still viewed as an important target for treating essential tremor (Blomstedt et al., 2007; Pahwa et al., 2006) and renewed interest has been shown in the targeting of the globus pallidus as a stimulation point for treating dystonia (Houeto et al., 2007; Vidailhet et al., 2007). We also assume that anatomically based inferences of surgical targets cannot be properly made if the gross anatomy of the atlas does not properly match the patient.

All raters were trained according to rules developed by the authors. The medial and superior borders of the caudate were defined in relation to the lateral ventricle, laterally by the internal capsule, and inferiorly by the external capsule. The putamen was defined medially by the internal capsule and the globus pallidus, laterally by the external capsule, superiorly by the internal capsule and corona radiata, and inferiorly by the external capsule. Raters were asked to join the two structures when possible. The thalamus was defined medially by the third ventricle wall, laterally by the internal capsule, anteriorly by the globus pallidus and the internal capsule, posteriorly by the extents of the pulvinar, superiorly by the lateral ventricle, and inferiorly by the inferior horn of the lateral ventricle. The globus pallidus was defined anteriorly, medially, and superiorly by the internal capsule, laterally by the putamen, posteriorly by the thalamus and the internal capsule, and inferiorly by the inferior horn of the lateral ventricle. To improve their initial labeling estimates, all raters were asked to create surface renderings (Lorensen and Cline, 1987) to better visualize the 3D consistency of their labels. If a series of voxels were found to be labeled that did not seem to coherently create a 3D anatomical structure, the raters were asked to revise their labels accordingly.

As shown in Fig. 1, patient scans acquired with the head-frame in a body coil suffer from a lack of contrast and resolution, therefore making it difficult for the raters to properly label the sub-cortical nuclei. While the stereotaxic head-frame provides an coordinate system which is used to guide surgical procedures, contrast of the subcortical nuclei is typically sacrificed. This trade-off mimics the typical clinical situation, where the accurate localization of surgical tools is gained at the expense of the visualization of subcortical nuclei granted by an MRI scan.

To aid the raters, the pre-operative scan used for diagnostic purposes was registered to the head-frame scan (using a rigid-body transformation (Collins et al., 1994)) and then the two were averaged to improve the signal-to-noise ratio and contrast to facilitate the identification of the sub-cortical nuclei mentioned above (see Fig. 5). Although this gives the raters a clear advantage over the

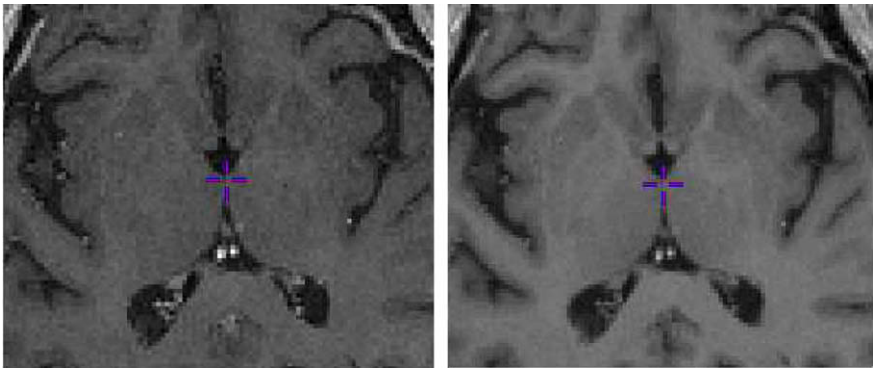

Fig. 5. Example of data used for validation of warping techniques. (From left to right) Axial sections MRI volumes acquired with head-frame, and the average of a volume acquired with a head-frame and without a head-frame. Note the difference between resolution and contrast of the two volumes. Averaging the two volumes increases the ease for raters to properly identify the subcortical nuclei being assessed in Section 3.2.

automated atlas warping techniques, it was done to facilitate the accurate identification of subcortical structures.

3.2.1. Derivation of a silver standard

To confidently use the manually labeled data to assess the atlas-to-template and atlas-to-patient warping techniques, the consistency of manual raters was verified. In the absence of an anatomical "gold standard" a series of "silver standards" were created from the manual labels in a leave-one-out fashion, where a single rater's labels are compared to a silver-standard developed through the agreement of other labels. Five different silver standards were created for each structure identified on each subject. This is achieved by sequentially eliminating each raters labels from the silver standard creation. After the elimination of a single rater, the remaining voxels in the silver standard are set to one if at least three of the four raters have labeled the voxel. In this fashion, each rater can be evaluated against the silver standard created independently of their own labels, thus an unbiased estimate of the raters performance can be achieved. Each silver standard created was also evaluated against the automated warping techniques, allowing the calculation of five different comparisons for each structure on each subject.

3.2.2. The kappa overlap metric

The labels defined by the atlas or a single rater (defined as the test structure) are compared to a silver standard by determining the level of overlap using the kappa metric (κ):

$$\kappa = \frac{2a}{(2a+b+c)}, \tag{2}$$

where a is the number of voxels common to the test structure and the silver standard, and $b+c$ represents the sum of the voxels uniquely identified by either the test structure or the silver standard, respectively.

The kappa score is extremely sensitive to differences in two label sets. The simulation performed in Fig. 6 shows a coronal slice through the striatum from the atlas. If the striatum is translated by 0.5 mm in all three dimensions, the kappa score decreases from 1.0 (for the label on itself) to 0.87. If the striatum if translated by 1 mm in all three dimension, then the kappa score decreases further to 0.80. A dilation of 1 mm causes the kappa score to decrease to 0.92, whereas an erosion of 1 mm will cause the kappa score to decrease to 0.91. Typically, scores greater than 0.7 are deemed acceptable in the segmentation and classification literature.

3.2.3. Evaluation of quality of manual labeling

To ensure that the manually labeled data were produced in a consistent manner, all of the manual labels for each rater were first evaluated against the silver standard. Statistical differences between the kappas measured for the labels of each rater were assessed using a repeated measures multiple analysis of variance (MANOVA) where the kappas of each rater were considered the main effect and hemisphere and structure were used as covariates.

3.2.4. Anatomical validation of atlas-to-template warping technique

In order to ensure that the atlas-to-patient warping techniques are properly validated the initial atlas-to-template warp must first be evaluated. This will assess the registration error that would be propagated in Methods A and B. The validation of the initial atlas-to-template warp discussed in Section 2.1 is assessed using the leave-one-out technique. Since the Colin27 MRI template is a

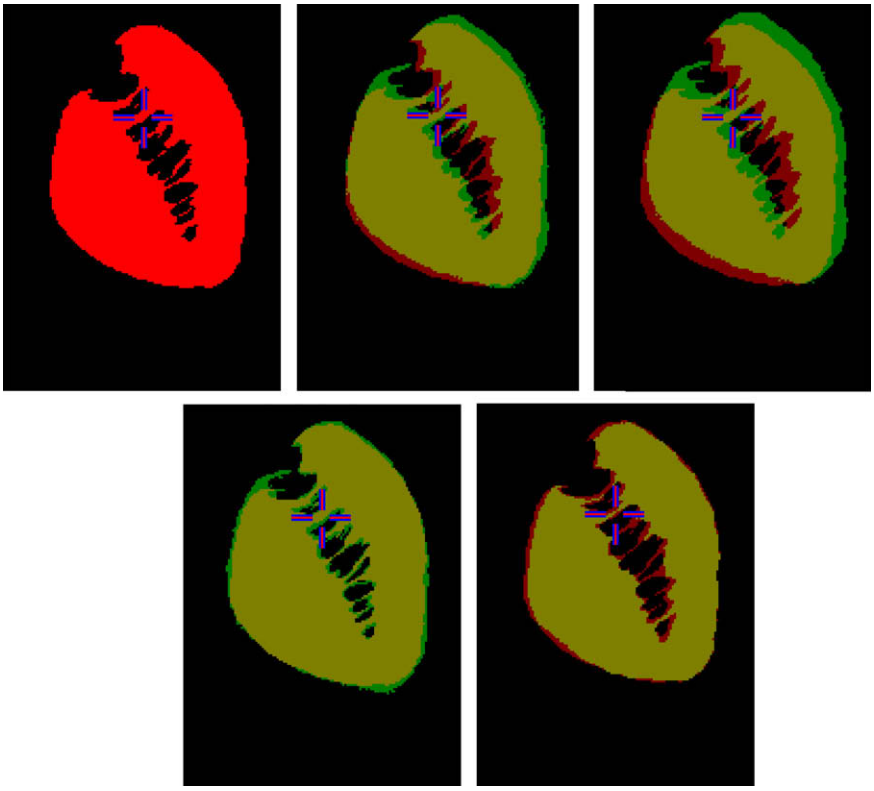

Fig. 6. Changing kappa values for a wandering striatum. (Top from left to right) Coronal slice through the original striatum defined by the atlas, the striatum translated by 0.5 mm in all three dimension ($\kappa = 0.87$), and striatum displaced by 1 mm in all three dimensions ($\kappa = 0.80$). (Bottom from left to right) Striatum is dilated by 1 mm voxel in all three dimensions ($\kappa = 0.92$) and is eroded by 1 mm in all three dimensions ($\kappa = 0.91$). In all images, the label in red represents the original striatum, and the label in green represents the simulated change in the striatum, and yellow represents the overlap between the two. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

high resolution template and each of the five raters performed only one segmentation of this template, a mean kappa is calculated for all the raters. Results for each structure were compared using a pooled *t*-test to assess statistical differences.

3.2.5. Anatomical validation of atlas-to-patient warping techniques

The overlap of each the silver standards with the labels derived from the five manual raters and the three atlas-to-patient warping techniques were analyzed using a repeated measures MANOVA. The kappas values from each technique (manual and automatic) was considered as the main effect and the hemispheres and structures were treated as covariates. In cases where statistical differences were found, the nature of these differences were analyzed with a post hoc Tukey HSD test. This test will group results which show no significant differences. In the results section these groups are identified using as groups **a**, **b**, **c**, etc. Methods or raters identified by group **a** have the highest kappa, and no significant differences between their results. Similarly, methods or raters identified in group **b** have mean kappas significantly lower than methods and raters in group **a**, but show no statistically significant differences between themselves.

3.3. Electro-physiological validation

The functional validation uses monopolar stimulation of responses to identify the sensory cutaneous thalamus (ventralis caudalis (Vc) (Hassler and Riechert, 1954), ventral posterior (VP) (Hirai and Jones, 1989)). Stimulations were done using a curved retractable monopolar stimulation electrode (Atkinson et al., 2002; Duerden et al., 2003; St-Jean et al., 1998). Cutaneous paresthesias on the face, arm, hemi-body or leg were evoked using difference levels of electrical stimulation (0.25–1 V; at 185 Hz with 2 ms pulse duration). Points at which paresthesias were evoked at a threshold of 0.5 V are used in the present study, localizing the hand area of the sensory cutaneous thalamus (VP, lateral part, VPL, to within 2 mm (Atkinson et al., 2002)). Using surgical guidance software (St-Jean et al., 1998), the lesion target location, the extension of the electrical stimulator, and angles of declination and azimuth the location of the tip of the probe can be identified in its frame-space coordinates of the tip of stimulator can be found. For all atlas-to-subject warping techniques, the locations of stimulation points were transformed onto the Colin27 MRI average or the pseudo-MRI using the inverse atlas-to-subject transformation. Only stimulations that evoked sensory paresthesias were used for this validation, yielding a total of 61 functional data points from the nine patients. Three different accuracy measures were defined:

1. The transformed functional data points, were analyzed by determining how many fell inside the 3D boundary defined by the sensory thalamus (see Fig. 7).
2. The distance between the stimulation point and the border of the sensory thalamus defines the second metric. It was estimated using the chamfer distance function (Borgefors, 1984), an approximation of the Euclidean distance. In this metric, points inside the sensory thalamus are considered positive and points falling outside the sensory thalamus are considered negative. An ANOVA was used to assess if there were any differences between groups.
3. The third metric is also distance based. It is equal to the Euclidean distance between each point and the center of mass of the sensory thalamus. An ANOVA was used to assess if there were differences between groups.

4. Results

4.1. Quality of manual labeling

The results of the manual labeling are given in Table 3. A MANOVA using the kappas of the raters as the main effect and structure and hemisphere as covariates showed no significant effect due to hemisphere ($p = 0.5591$). Subsequently the covariate of hemisphere was eliminated from the analysis. An effect of structure ($F = 106.9225$, $p < 0.0001$, DF = 2) and rater ($F = 30.4028$,

Table 3
Result from post hoc Tukey-Kramer HSD test for kappas from manual raters

	Groups	Mean ± SD	Range
Globus Pallidus			
Rater 1	b	0.530 ± 0.106	0.322–0.732
Rater 2	a	0.687 ± 0.096	0.441–0.839
Rater 3	c	0.431 ± 0.066	0.308–0.597
Rater 4	a	0.641 ± 0.109	0.429–0.835
Rater 5	a	0.626 ± 0.112	0.374–0.809
Striatum			
Rater 1	b	0.799 ± 0.041	0.723–0.863
Rater 2	a	0.838 ± 0.026	0.795–0.922
Rater 3	c	0.763 ± 0.033	0.693–0.820
Rater 4	a	0.849 ± 0.033	0.763–0.895
Rater 5	a, b	0.825 ± 0.036	0.752–0.895
Thalamus			
Rater 1	a	0.825 ± 0.076	0.567–0.885
Rater 2	a	0.836 ± 0.099	0.454–0.907
Rater 3	a	0.823 ± 0.026	0.774–0.866
Rater 4	a	0.794 ± 0.055	0.676–0.872
Rater 5	a	0.843 ± 0.034	0.763–0.886

Test was performed over kappas measurements for each structure with respect to silver standard. Raters are grouped based on the result of Tukey-Kramer HSD post hoc analysis ($p < 0.05$). The mean minimum, maximum, and standard deviations of the kappa's are also provided.

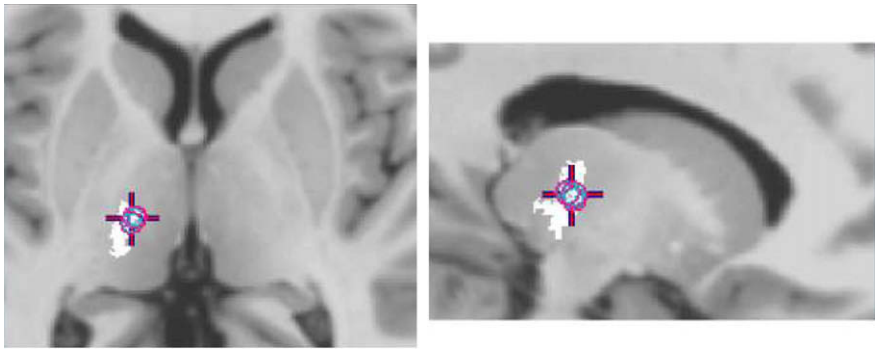

Fig. 7. Example of functional validation of the sensory thalamus. Landmarks represent a point where sensory responses were invoked using intra-operative electrophysiological confirmation. The white label represents the sensory thalamus as defined by the atlas.

$p < 0.0001$, DF = 4), as well as interactions between structure and rater were observed ($F = 14.0902$, $p < 0.001$, DF = 9). A post hoc test (Tukey-Kramer HSD) assessing how each structure was labeled by the raters was also performed. This experiment revealed, that while all raters agreed on the definition of the thalamus ($\bar{\kappa}$ range: 0.794–0.843), only three agreed on the definition of the globus pallidus ($\bar{\kappa}$ range: 0.431–0.687) and the striatum ($\bar{\kappa}$ range: 0.763–0.849).

The low kappa values seen for the globus pallidus, and the inter rater variability for the globus pallidus and the thalamus were subsequently investigated. Our observations showed that two main causes attributed to the variability and the low kappa values. The first is the intra-rater disagreement of the manual raters with respect to the location of the posterior end of the globus pallidus. The second was the longer definition of the globus pallidus in the atlas in comparison to the definition provided by the manual raters.

4.2. Anatomical atlas-to-template warping validation

Table 4 shows the mean kappa, standard deviation, and range of the kappas for all structures and each structure individually, when the labels from the manual raters and the warped atlas labels are compared to the silver standard. The results of the pooled *t*-tests for all structures (DF = 53) and each individual structure (DF = 18) are also shown in Table 4. No significant differences are seen in the definitions for all structures or the globus pallidus defined by the manual raters ($\bar{\kappa}$ for all structures = 0.86, $\bar{\kappa}$ for globus pallidus = 0.76) and those of the warped atlas ($\bar{\kappa}$ for all structures = 0.83, $\bar{\kappa}$ for globus pallidus = 0.73).

Although statistically significant differences are seen in the striatum, the mean kappas values are similar for the manual raters ($\bar{\kappa} = 0.90$) and the warped atlas ($\bar{\kappa} = 0.89$). The warped atlas also shows a much tighter range of kappa values ($\bar{\kappa}$: 0.88—0.89) than the manual raters ($\bar{\kappa}$: 0.87—0.92). Statistical differences are also seen for the definition of the thalamus, while the mean kappa of the warped atlas is lower ($\bar{\kappa} = 0.86$), than that of the manual raters ($\bar{\kappa} = 0.90$), the range of values is much tighter for the warped atlas ($\bar{\kappa}$: 0.84—0.86), than the manual raters ($\bar{\kappa}$: 0.87—0.93).

4.3. Anatomical atlas-to-patient warping validation

For the automated method validation, for each structure, Method B shows the highest mean kappa value over all structures. All automated warping techniques perform better than Rater 3. For the globus pallidus, there are no statistically significant differences between the automated warping methods and Raters 1 and 2. The other raters have significantly higher kappa values than the automated raters.

The mean, standard deviation, and range of the data for anatomical validation for the atlas-to-patient warping techniques using clinical data are given in Tables 5–7. No effect from hemisphere was seen ($F = 0.4450$, $p = 0.5079$, DF = 1), thus it was eliminated from the analysis and only structure was kept as a covariate. A MANOVA found the affect of structure and method to be significant ($F = 117.1950$, $p < 0.0001$, DF = 7). Differences between the kappas of each method were analyzed by structure using a Tukey-Kramer HSD post hoc test (see also Tables 5–7). For the globus pallidus and the striatum, the manual raters have highest mean kappas. Method B shows highest mean kappas for the globus pallidus ($\bar{\kappa} = 0.549$) and the striatum ($\bar{\kappa} = 0.750$). In the striatum, Methods A and B performed statistically better than Method C. No statistical differences were found between the manner in which the thalamus was identified by Method B ($\bar{\kappa} = 0.818$) or any of the manual raters. However, Methods A ($\bar{\kappa} = 0.787$) and C ($\bar{\kappa} = 0.762$) show acceptable mean kappas for the thalamus as well.

4.4. Electro-physiological validation

The results of the functional validation show that 39 of the stimulations points recorded (64%) fell into the spatial limits of the sensory thalamus for both the automatic warping Methods B and C, compared to 35 points (57%) for Method A.

The ANOVA performed on the data from chamfer distance from the border of the sensory thalamus test also yielded no significant differences between methods ($F = 1.1574$, $p < 0.3175$, DF = 2) (see Table 8).

Using an ANOVA, no significant differences in distances were found between each functional data point and the center of mass of the sensory thalamus were found between any of the methods ($F = 0.5897$, $p < 0.5555$, DF = 2) (see Table 9).

Table 4
Results of atlas-to-template warping: kappa values for the striatum, globus pallidus, and thalamus of the Colin27 template

Structure	Manual rater $\bar{\kappa} \pm$ SD (range)	automatic warping $\bar{\kappa} \pm$ SD (range)	*t*-value ($p < t$)
All	0.86 ± 0.068 (0.65–0.93)	0.83 ± 0.055 (0.64–0.90)	1.86 (0.192)
Globus pallidus	0.79 ± 0.075 (0.65–0.88)	0.76 ± 0.013 (0.74–0.78)	1.31 (0.103)
Striatum	0.90 ± 0.017 (0.87–0.92)	0.89 ± 0.004 (0.88–0.89)	2.43 (0.013)
Thalamus	0.90 ± 0.018 (0.87–0.93)	0.86 ± 0.007 (0.84–0.86)	7.34 (0.0001)

Mean, standard deviation, and range values for manual raters and the pseudo-MRI based atlas-to-template warping techniques are given.

Table 5
Result from post hoc Tukey-Kramer HSD test for kappas from manual raters and automatic methods for the globus pallidus

Globus pallidus	Groups	Mean ± SD	Range
Rater 1	b, c, d	0.530 ± 0.106	0.322–0.732
Rater 2	a	0.687 ± 0.096	0.441–0.839
Rater 3	d	0.431 ± 0.066	0.308–0.597
Rater 4	a	0.641 ± 0.109	0.429–0.835
Rater 5	a, b	0.626 ± 0.112	0.374–0.809
Method A	d	0.522 ± 0.105	0.305–0.702
Method B	d	0.549 ± 0.107	0.319–0.745
Method C	d	0.509 ± 0.075	0.358–0.690

Test was performed over kappa measurements for the globus pallidus. Raters are grouped based on the result of Tukey-Kramer HSD post hoc analysis ($p < 0.05$). The mean minimum, maximum, and standard deviations of the kappa's are also provided.

Table 6
Result from post hoc Tukey-Kramer HSD test for kappas from manual raters for the striatum

Striatum	Groups	Mean ± SD	Range
Rater 1	a, b	0.799 ± 0.041	0.723–0.863
Rater 2	a	0.838 ± 0.026	0.795–0.922
Rater 3	b, c	0.763 ± 0.033	0.693–0.820
Rater 4	a	0.849 ± 0.033	0.763–0.895
Rater 5	a, b	0.825 ± 0.036	0.752–0.895
Method A	c	0.722 ± 0.065	0.570–0.842
Method B	c	0.750 ± 0.059	0.603–0.858
Method C	d	0.648 ± 0.082	0.506–0.810

Test was performed over kappa measurements for the globus pallidus. Techniques are grouped based on the result of a Tukey-Kramer HSD post hoc analysis ($p < 0.05$). The mean minimum, maximum, and standard deviations of the kappa's are also provided.

Table 7
Result from post hoc Tukey-Kramer HSD test for kappas from manual raters for the thalamus

Thalamus	Groups	Mean ± SD	Range
Rater 1	a, b	0.825 ± 0.076	0.567–0.885
Rater 2	a	0.836 ± 0.099	0.454–0.907
Rater 3	a, b	0.823 ± 0.026	0.774–0.866
Rater 4	a, b, c	0.794 ± 0.055	0.676–0.872
Rater 5	a	0.843 ± 0.034	0.763–0.886
Method A	b	0.787 ± 0.041	0.693–0.846
Method B	a, b	0.818 ± 0.036	0.690–0.876
Method C	c	0.762 ± 0.058	0.653–0.851

Test was performed over kappa measurements for the thalamus. Raters are grouped based on the result of Tukey-Kramer HSD post hoc analysis ($p < 0.05$). The mean minimum, maximum, and standard deviations of the kappa's are also provided.

Table 8
Chamfer distance analysis of each of functional stimulation points from the border of the sensory thalamus

Method	Mean (mm)	SD (mm)	Range (mm)
Method A	−1.2	1.2	−3.74 to 0.94
Method B	−1.3	1.3	−4.17 to 1.15
Method C	−1.0	1.8	−5.19 to 1.55

Test was performed over distance estimated from the edge of the structure. An ANOVA revealed no significant differences. The mean, standard deviations, and range of values of the chamfer distance are also provided.

Table 9
Distance analysis of each of functional stimulation points from the center of mass (COM) of the sensory thalamus

Method	Mean (mm)	SD (mm)	Range (mm)
Method A	6.6	2.8	1.3–11.2
Method B	7.2	2.4	2.8–10.8
Method C	5.8	2.6	1.3–10.2

Test was performed over distance estimates for each data point from the COM of the structure. An ANOVA revealed no significant difference. The mean, standard deviations, and range of values of the Euclidean distance are also provided.

5. Discussions

5.1. Summary

In this paper we developed techniques for the anatomical and electro-physiological validation of atlas-to-template and atlas-to-patient warping techniques using a digital atlas of the basal ganglia and thalamus previously developed in our group (Chakravarty et al., 2006a). The atlas was nonlinearly warped using the ANIMAL algorithm (Collins and Evans, 1997; Collins et al., 1995) to a high resolution MRI template (Holmes et al., 1998) via a direct atlas-to-template warping scheme that relies upon a pseudo-MRI representation of the atlas (Chakravarty et al., 2005, 2006a).

Three different atlas-to-patient warping procedures were validated using anatomical and electro-physiological techniques. Method A uses the Colin27 MRI template and produced a nonlinear transformation defined by vectors spaced on a grid with 2 mm spacing. Method B also uses the Colin27 MRI template for atlas-to-patient registration, however the final nonlinear transformation is defined on a grid with 1 mm spacing. Method C relies directly on the pseudo-MRI to estimate the atlas-to-patient transformation and output a final grid using the same ANIMAL parameters of Method B.

Anatomical and electro-physiological criteria were used for validation of all warping methods. Five different manual raters defined anatomical silver standard using a "leave-one-out" technique. The striatum, globus pallidus, and thalamus of the Colin27 MRI template and nine patients were manually labeled by each rater. The overlap between the automatic methods and the manual raters were evaluated against each silver standard. Electro-physiological evaluation was done using macro-stimulation points used to define the sensory thalamus intra-operatively. All points were warped back to the Colin27 template using all warping method. The number of stimulations falling within the atlas definition of the sensory thalamus, the Euclidean distance of each point to the center of mass of the atlas sensory thalamus, and the chamfer distance from the edge of the sensory thalamus were used as metrics to evaluate the accuracy of the warping.

5.2. Quality of manual labeling

Inspection of the quality of manual rating (Table 3) demonstrates significant differences in the level of agreement between raters and the silver standards of the globus pallidus and the striatum. In addition, the mean kappa values for any of the raters failed to reach what are considered an "acceptable" level of agreement ($\bar{\kappa} > 0.7$) for the globus pallidus. In addition, large standard deviations and a wide range of values are observed for all raters. These large levels of disagreement of the definition of the striatum and globus pallidus occur at the posterior end of the structures where these structures taper and grey and white matter contrast become more difficult to distinguish.

In the case of the globus pallidus there are further confounding factors such as low contrast and resolution which make the consistent manual identification of this structure difficult. Furthermore, lower kappas can be expected in smaller structures. Structure which have a high surface area-to-volume ratio are additionally penalized by the kappa metric. However the main source of inter-rater disagreement for this structure is due to Raters 2 and 3. This further demonstrates the difficulty of using manual raters for target identification as inter-rater reliability may be low even when using qualified trained manual raters. While five raters used in this study have expert knowledge of the subcortical anatomy, their personal definition of the structures may be slightly different even if they were in accordance with the labeling rules used. These differences may still cause the large variations seen in the results. Atlas-based identification minimizes the subjectivity of the structure identification inherent to manual labeling. The larger agreement in the definition of the thalamus is likely due to its more morphologically homogeneous structure.

5.3. Validation of the atlas-to-template warp

Analysis of the initial atlas-to-template warp shows acceptable levels of agreement ($\bar{\kappa} > 0.7$) for all structures defined by the atlas. While the thalamus kappas of the automated technique are statistically different than the manual raters, a much smaller range of values and standard deviations are observed for all structures when compared to the manual raters. This implies that the atlas is more consistently in agreement with the silver standard and that there is a larger inter-rater variability.

5.4. Anatomical validation of atlas-to-patient warping techniques

In the case of the striatum, Method C does not achieve acceptable levels of agreement and is statistically worse than all raters and techniques. Methods A and B show statistically significant differences with all raters except for Rater 3.

No statistically significant differences are found between the thalami defined by Method B and the manual raters. Method C performs significantly worse. Method A also shows similarity to three of the raters and Method B. Since Method B is always better than

the other warping techniques, we recommend that it be used for atlas-to-patient warping, despite the expense of computational time.

For all three structures the variability of Methods A and B are in line with that of the manual raters. Method C does not perform as well in the manner in which it defines the three subcortical anatomical regions. While the pseudo-MRI has shown to be an excellent method to use for warping the atlas to the template volume, it may not adequately resemble the T1-weighted MRI data typically used in surgical planning, revealing that our initial hypothesis was incorrect. Firstly the pseudo-MRI was developed to match the contrast and intensities of the Colin27 template. This template, achieves a significant reduction in noise due to the nature of its derivation. Furthermore, each structure of the pseudo-MRI is homogenous throughout, thus the dynamic range is limited by the number of structures present in the atlas. This is sufficient when attempting to match the pseudo-MRI to a volume such as the Colin27 template, however it may have limited the accuracy of the atlas-to-patient transformation estimation. The results for the atlas-to-template warping suggest that this initial warp is quite accurate when compared to the silver standards. Furthermore, the atlas is limited to structures in the basal ganglia. This does not aid the use of the ANIMAL algorithm, as it depends on the fitting of structures not just in the basal ganglia, but also in its immediate region to provide the optimal fit. As a result the whole brain and ROI based strategies of Methods A and B are benefitted.

The largest variations were observed in the labeling of the striatum and the globus pallidus. The resolution of the serial histological sections from which the atlas was derived allows the posterior portions of these structures to be well identified in the atlas. However the limited resolution of the pre-operative MRI does not allow this. This demonstrates one of the main differences between the atlas and the labels derived by the manual raters, as the identification of structures which are difficult to resolve using standard pre-operative MRI data were more easily resolved on the atlas.

The detailed analysis of these structures is done for two main the reasons. The first is the typical assumption made in the pre-operative planning phase is that functional somatopy can be accounted for by proper labeling of the anatomy (Ganser et al., 2004; Penfield and Boldrey, 1937; Yelnik et al., 2007). The second reason is that we required a validation technique that best mimics the clinical situation where a single rater (a technician or a surgeon) is typically used to manually label the subcortical structures to aid intra-operative visualization.

One of the main novelties of this study resides in the use and derivation of a silver standard from manual raters for the comparison of different registration techniques used for atlas warping. The results from the verification of the quality of the manual raters demonstrate variability in the rater definitions of all structures and underscores the difficulty of using these definitions of the anatomy for comparison. Work by Warfield et al. (2004) has addressed this issue by developing an expectation maximization algorithm which estimates an optimal segmentation from the different segmentation methods being evaluated. Using this technique, each method can then be weighted depending upon its estimated performance level with respect to the other methods being tested. However, if the segmentations used do not include a silver or gold standard, then the results obtained using this technique will assess the consistency of the segmentation methods, and not necessarily their accuracy. Despite the variability of the manual raters definitions, each rater used is known to have an accurate definition of the structures tested and the method used in this study shows that the raters can effectively provide an upper limit for the accuracy of anatomical labeling.

5.5. Functional validation of atlas-to-patient warping techniques and comparison to other point-based techniques

None of the three metrics based on electrophysiological recordings demonstrate a statistically significant difference in which these techniques localize, harder to define, anatomical structures. One aspect of this validation that will significantly affect the results is the unknown nature of the point spread function. This has been a common problem amongst groups seeking validation of their atlas-warping results. For example, in the work of Finnis et al. (2003), positive stimulations are coded with a sphere proportional to the strength of the stimulation, however it is has never been proven that this is an accurate method for the representation of the actual stimulation. In one aspect of their functional validation, 71% of microelectrode stimulations evoking a visual response (used to determine pallidal targets) were compared with their location relative to the optic track. Since the microelectrodes likely have a smaller point spread function, more accurate results can be expected. Similarly, the results from the extension of this original work from Guo et al. (2006) demonstrate that the mean difference between their atlas predicted surgical targets is on the order of 2.18–2.33 mm for thalamotomy, pallidotomy, thalamic DBS, and STN DBS.

Other groups who have tried to predict the final target location have shown errors of greater than a millimeter in their target localization. For example in Nowinski et al. (2000) targeting errors can be reduced from 3.07 to 1.44 mm for thalamotomy, pallidotomy, and pallidal DBS. In recent work in STN DBS targeting by Sanchez Castro et al. (2006b) demonstrate approximately 1.72 and 1.77 mm difference between their expected surgical targets and those targets by warping their atlas using B-splines and the Demons algorithm, respectively. Similarly D'Haese et al. show that their system is able to predict the final STN DBS location to within 2 mm (D'Haese et al., 2005a).

What can be elucidated from previous studies is that anatomical and functional atlases can localize targets to within 1.5–2.0 mm, thereby revealing the limitations of most atlas warping and atlas construction techniques. Furthermore, in a previous study from our group (Atkinson et al., 2002), demonstrated that the stimulations locating the surgical target be accurate to within 2 mm when compared with stimulations used to confirm tremor arrest.

The above interpretations support the results of our functional validation as the atlas-based and electrophysiological factors may compound errors from the our analysis. The results demonstrating that the majority of all points used in this validation process overlapped with the atlas definition of the sensory thalamus is very encouraging. For the experiment using the chamfer distance based metric we see that the mean distance for all techniques is on the order of 1–1.5 mm outside the border and is within the accuracy of macroelectrode stimulator. Furthermore, the point spread function may indeed be larger than the voxel size of the data being used. Nonetheless, it is a good indication of target localization capabilities of the current atlas-to-patient warping techniques.

6. Conclusions and future work

Based on the findings of this study it can be inferred that Method B is the most accurate method, of the three methods tested, to customize subcortical structures to a particular patients anatomy. The results are comparable to the variability of the manual raters. However this technique does not behave in a manner which is statistically better than Method A, therefore it is left to the user to decide between the trade-off for accuracy and time taken to develop the high-resolution vector field defined by Method B. While the

direct, pseudo MRI based method does an acceptable job with the identification of the thalamus it is not a suitable method for estimating atlas-to-patient transformations; however, we have shown that it is an excellent method with which to define the initial atlas-to-template warp. Subsequently, the atlas-to-template warping errors which are propagated through the use of the template based Methods A and B yield acceptable results when evaluating the sub-cortical anatomy.

Future work will include applying these techniques to expand the functional atlas we have begun to build (Chakravarty et al., 2006b). We will also expand the validation techniques discussed to perform a broader comparison of standard and nonlinear atlas-to-patient warping techniques.

References

Atkinson, J., Collins, D.L., Bertrand, G., Peters, T.M., Pike, B., Sadikot, A.F., 2002. Optimal location of thalamotomy lesions for tremor associated with Parkinson's disease: a probabilistic analysis based on postoperative magnetic resonance imaging and integrated digital atlas. Journal of Neurosurgery 96 (5), 854–866.

Bardinet, E., Dormont, D., Malandain, G., Bhattacharjee, M., Pidoux, B., Saleh, C., Cornu, P., Ayache, N., Agid, Y., Yelnik, J., 2005. Retrospective cross-evaluation of a histological and deformable 3D atlas of the basal ganglia on series on Parkinsonian patients treated by deep brain stimulation. In: Eighth International Conference on Medical Image Computing and Computer Assisted Intervention MICCAI 2005, Palm Springs, USA, October. Lecture Notes in Computer Science, vol. 2. Springer, pp. 385–393.

Behrens, T.E.J., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshot, C.A.M., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M., 2003. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nature Neuroscience 6 (7), 750–757.

Benabid, A.L., Koudsié, A., Benazzouz, A., Le Bas, J.F., Pollak, P., 2002. Imaging of subthalamic nucleus and ventralis intermedius of the thalamus. Movement Disorders 17 (S3), 123–129.

Bertrand, G., 1982. Computers in functional neurosurgery. In: Schaltenbrand, G., Walker, A.E. (Eds.), Stereotaxy of the Human Brain: Anatomical, Physiological and Clinical Applications. Thieme, New York.

Bertrand, G., Olivier, A., Thompson, C.J., 1973. The computerized brain atlas: its use in stereotaxic surgery. Transactions of the American Neurological Association 98, 233–237.

Bertrand, G., Olivier, A., Thompson, C.J., 1974. Computer display of stereotaxic brain maps and probe tracts. Acta Neurochirurgica 21 (Suppl.), 235–243.

Blomstedt, P., Hariz, G.M., Hariz, M.I., Koskinen, L.-O.D., 2007. Thalamic deep brain stimulation in the treatment of essential tremor: a long-term follow-up. British Journal of Neurosurgery 21 (5), 504–509.

Bookstein, F.L., 1989. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (6), 567–585.

Borgefors, G., 1984. Distance transformations in arbitrary dimensions. Computer Vision, Graphics, and Image Processing 27 (3), 321–345.

Chakravarty, M.M., Sadikot, A.F., Germann, J., Bertrand, G., Collins, D.L., 2005. Anatomical and electrophysiological validation of an atlas for neurosurgical planning. In: Eighth International Conference on Medical Image Computing and Computer Assisted Intervention MICCAI 2005, Palm Springs, USA, October. Lecture Notes in Computer Science, vol. 2. Springer, pp. 394–401.

Chakravarty, M.M., Bertrand, G., Hodge, C.P., Sadikot, A.F., Collins, D.L., 2006a. The creation of a brain atlas for image guided neurosurgery using serial histological data. NeuroImage 30 (2), 359–376.

Chakravarty, M.M., Sadikot, A.F., Mongia, S., Bertrand, G., Collins, D.L., 2006b. Towards a multi-modal atlas for neurosurgical planning. In: Ninth International Conference on Medical Image Computing and Computer Assisted Intervention MICCAI 2006, Copenhagen, Denmark, October. Lecture Notes in Computer Science, vol. 2. Springer, pp. 389–396.

Cohn, M.C., Hudgins, P.A., Sheppard, S.K., Starr, P.A., Bakay, R.A.E., 1998. Pre- and postoperative evaluation of stereotactic pallidotomy. American Journal of Neuroradiology 19 (6), 1075–1080.

Collins, D.L., Evans, A.C., 1997. ANIMAL: validation and application of non-linear registration-based segmentation. International Journal of Pattern Recognition and Artificial Intelligence 11 (8), 1271–1294.

Collins, D.L., Holmes, C.J., Peters, T., Evans, A.C., 1995. Automatic 3-D model based neuroanatomical segmentation. Human Brain Mapping 3 (3), 190–208.

Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. Journal of Computer Assisted Tomography 18 (2), 192–205.

Deoni, S.C.L., Peters, T.M., Rutt, B.K., 2005. High resolution T1 and T2 mapping of the brain in clinically acceptable time with DESPOT1 and DESPOT2. Magnetic Resonance in Medicine 53 (1), 237–241.

D'Haese, P.F., Cetinkaya, E., Konrad, P.E., Kao, C., Dawant, B.M., 2005a. Computer-aided placement of deep brain stimulators: from planning to intraoperative guidance. IEEE Transactions on Medical Imaging 24 (11), 1469–1478.

D'Haese, P.F., Pallavaram, S., Niermann, K., Spooner, J., Kao, C., Konrad, P.E., Dawant, B.M., 2005b. Automatic selection of DBS target points using multiple electrophysiological atlases. In: Eighth International Conference on Medical Image Computing and Computer Assisted Intervention MICCAI 2005, Palm Springs, USA, October. Lecture Notes in Computer Science, vol. 2. Springer, pp. 427–434.

Duerden, E.G., Finnis, K.W., Peters, T.M., Sadikot, A.F., 2003. A method for analysis of electrophysiological responses obtained from the motor fibers of the human internal capsule. In: Fifth International Conference on Medical Image Computing and Computer Assisted Intervention MICCAI 2003, Montreal, Canada, November. Lecture Notes in Computer Science. Springer, pp. 50–57.

Duval, C., Strafella, A.P., Sadikot, A.F., 2005. The impact of ventrolateral thalamotomy on high-frequency components of tremor. Clinical Neurophysiology 116 (6), 1391–1399.

Eskandar, E.N., Cosgrove, G.R., Shinobu, L.A., 2001. Surgical treatment of parkinson's disease. Journal of American Medicine 286 (24), 3056–3059.

Finnis, K.W., Starrveld, Y.P., Parrent, A.G., Sadikot, A.F., Peters, T.M., 2003. Three-dimensional database of subcortical electrophysiology for image-guided stereotactic functional neurosurgery. IEEE Transactions on Medical Imaging 22 (1), 93–104.

Ganser, K.A., Dickhaus, H., Metzner, R., Wirtz, C.R., 2004. A deformable digital brain atlas system according to Talairach and Tournoux. Medical Image Analysis 8 (1), 3–22.

Gloor, P., 1997. The Temporal Lobe and Limbic System. Oxford University Press, New York.

Gross, R.E., Lombardi, W.J., Lang, A.E., Duff, J., Hutchison, W.D., Saint-Cyr, J.A., Tasker, R.R., Lozano, A.M., 1999. Relationship of lesion location to clinical outcome following microelectrode-guided pallidotomy for Parkison's disease. Brain 122 (3), 405–416.

Guo, T., Finnis, K.W., Parrent, A.G., Peters, T.M., 2005. Development and application of functional databases for planning deep-brain neurosurgical procedures. In: Eighth International Conference on Medical Image Computing and Computer Assisted Intervention MICCAI 2005, Palm Springs, USA, October. Lecture Notes in Computer Science, vol. 1. Springer, pp. 835–842.

Guo, T., Finnis, K.W., Parrent, A.G., Peters, T.M., 2006. Visualization and navigation system development and application for stereotactic deep-brain neurosurgeries. Computer Aided Surgery 11 (5), 231–239.

Hardy, T.L., Bertrand, G., Thompson, C.J., 1979a. Organization and topography of sensory responses in the internal capsule and nucleus ventralis caudalis found during stereotactic surgery. Applied Neurophysiology 42, 335–351.

Hardy, T.L., Bertrand, G., Thompson, C.J., 1979b. The position and organization of motor fibers in the internal capsule found during stereotactic surgery. Applied Neurophysiology 42, 160–170.

Hardy, T.L., Bertrand, G., Thompson, C.J., 1979c. Position of the medial internal capsular border in relation to third-ventricular width. Applied Neurophysiology 42, 234–247.

Hardy, T.L., Bertrand, G., Thompson, C.J., 1979d. Thalamic recordings during stereotactic surgery. I. Surgery topography of evoked and nonevoked rhythmic cellular activity. Applied Neurophysiology 42, 185–197.

Hardy, T.L., Bertrand, G., Thompson, C.J., 1979e. Thalamic recordings during stereotactic surgery. II. Locations of quick-adapting touch-evoked (novelty) cellular responses. Applied Neurophysiology 42, 198–202.

Hardy, T.L., Bertrand, G., Thompson, C.J., 1980a. Position and organization of thalamic cellular activity during diencephalic recording. I. Pressure-evoked activity. Applied Neurophysiology 43, 18–27.

Hardy, T.L., Bertrand, G., Thompson, C.J., 1980b. Position and organization of thalamic cellular activity during diencephalic recording. II. Joint- and muscle-evoked activity. Applied Neurophysiology 43, 28–36.

Hardy, T.L., Bertrand, G., Thompson, C.J., 1980c. Topography of bilateral-movement-evoked thalamic cellular activity found during diencephalic recording. Applied Neurophysiology 43, 67–74.

Hardy, T.L., Bertrand, G., Thompson, C.J., 1981. Touch-evoked thalamic cellular activity. Applied Neurophysiology 44, 302–313.

Hassler, R., Riechert, T., 1954. Indikationen und lokalisations methode der gezielten hirnoperationen. Nervenarzt 25, 441.

Hellier, P., Barillot, C., Corouge, I., Gibaud, B., LeGoualhier, G., Collins, D.L., Evans, A.C., Malandain, G., Ayache, N., Christensen, G.E., Johnson, H.J., 2003. Restrospective evaluation of intersubject brain registration. IEEE Transactions on Medical Imaging 22 (9), 1120–1130.

Hirai, T., Jones, E.G., 1989. A new parcellation of the human thalamus on the basis of histochemical staining. Brain Research Reviews 14 (1), 1–34.

Holmes, C.J., Hoge, R., Collins, L., Woods, R., Toga, A.W., Evans, A.C., 1998. Enhancement of MR images using registration for signal averaging. Journal of Computer Assisted Tomography 22 (2), 324–333.

Houeto, J.L., Yelnik, J., Bardinet, E., Vercueil, L., Krystkowiak, P., Mesnage, V., Lagrange, C., Dormont, D., Le Bas, J.F., Pruvo, J.P., Tezenas du Moncel, S., Pollak, P., Agid, Y., Destée, A., Vidailhet, M.and French Simulation du Stimulation du Pallidum Interne dan la Dystonie Study Group, 2007. Acute deep-brain stimulation of the internal and external globus pallidus in primary dystonia: functional mapping of the pallidum. Archives of Neurology 64 (9), 1281–1286.

Johansen-Berg, H., Behrens, T.E.J., 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. Cerebral Cortex 15 (1), 31–39.

Krause, M., Fogel, W., Heck, A., Hacke, W., Bonsanto, M., Trenkwalder, C., Tronnier, V., 2001. Deep brain stimulation for the treatment of Parkinson's disease:

subthalamic nucleus versus globus pallidus internus. Journal of Neurology, Neurosurgery, and Psychiatry 70 (4), 464–470.
Lenz, F.A., Normand, S.L., Kwan, H.C., Andrews, D., Rowland, L.H., Jones, M.W., Seike, M., Lin, Y.C., Tasker, R.R., Dostrovsky, J.O., Lenz, Y.E., 1995. Statistical prediction of the optimal site for thalamotomy in parkinsonian tremor. Movement Disorders 10 (3), 318–328.
Lombardi, W.J., Gross, R.E., Trepanier, L.L., Lang, A.E., Lozano, A.M., Saint-Cyr, J.A., 2000. Relationship of lesion location to cognitive outcome following microelectrode-guided pallidotomy for Parkinson's disease: support for the existence of cognitive circuits in the human pallidum. Brain 123 (4), 746–758.
Lorensen, W.E., Cline, H.E., 1987. Marching cubes: a resolution 3D surface reconstruction algorithm. Computer Graphics 21 (4), 163–169.
Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997. Multimodality image registration by maximization of mutual information. IEEE Transactions on Medical Imaging 16 (2), 187–198.
Nowinski, W.L., Fang, A., Nguyen, B.T., Raphel, J.K., Jagannathan, L., Raghavan, R., Bryan, N.R., Miller, G.A., 1997. Multiple brain atlas database and atlas based neuroimaging system. Computer Aided Surgery 2 (1), 42–66.
Nowinski, W.L., Yang, G.L., Yeo, T.T., 2000. Computer aided stereotactic functional neurosurgery enhanced by the use of the multiple brain atlas database. IEEE Transactions on Medical Imaging 19 (1), 62–69.
Ono, M., Kubik, S., Abernathy, C.D., 1990. Atlas of the Cerebral Sulci. Georg Thieme Verlag/Thieme Medical Publishers, Stuttgart, Germany.
Otsuki, T., Jokura, H., Takahashi, K., Ishikawa, S., Yoshimoto, T., Kimura, M., Yoshida, R., Miyazawa, T., 1994. Stereotactic gamma-thalamotomy with a computerized brain atlas: technical case report. Neurosurgery 35 (4), 764–768.
Pahwa, R., Lyons, K.E., Wilkinson, S.B., Simpson Jr., R.K., Ondo, W.G., Tarsy, D., Norregaard, T., Hubble, J.P., Smith, D.A., Hauser, R.A., Jankovic, J., 2006. Long-term evaluation of deep brain stimulation of the thalamus. Journal of Neurosurgery 104 (4), 506–512.
Penfield, W., Boldrey, E., 1937. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain 60, 389–443.
Robbins, S., Evans, A.C., Collins, D.L., Whitesides, S., 2004. Tuning and comparing spatial normalization methods. Medical Image Analysis 8 (3), 311–323.
Rohde, G.K., Aldroubi, A., Dawant, B.M., 2003. The adaptive bases algorithm for intensity based non-rigid image registration. IEEE Transactions on Medical Imaging 22 (11), 1470–1479.
Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J., 1999. Nonrigid registration using free-form deformations: application to breast MR images. IEEE Transactions on Medical Imaging 18 (8), 712–721.
Sanchez Castro, F.J., Pollo, C., Cuisenaire, O., Villemure, J.G., Thiran, J.P., 2006a. Validation of experts versus atlas-based and automatic registration methods for subthalamic nucleus targeting on MRI. International Journal of Computer Assisted Radiology and Surgery 1 (1), 5–12.
Sanchez Castro, F.J., Pollo, C., Meuli, R., Maeder, P., Bach Cuadra, M., Cuisenaire, O., Villemure, J.G., Thiran, J.P., 2005. Cross validation of experts versus registration method for target localization in deep brain stimulation. In: Eighth International Conference on Medical Image Computing and Computer Assisted Intervention MICCAI 2005, Palm Springs, USA, October. Lecture Notes in Computer Science, vol. 1. Springer, pp. 417–424.
Sanchez Castro, F.J., Pollo, C., Mueli, R., Maeder, P., Cuisenaire, O., Bach Cuadra, M., Villemure, J.G., Thiran, J.P., 2006b. A cross validation study of deep brain stimulation targeting: from experts to atlas-based, segmentation-based, and automatic registration algorithms. IEEE Transactions on Medical Imaging 25 (11), 1440–1450.
Schaltenbrand, G., Bailey, W., 1959. Atlas of Stereotaxy of the Human Brain. Georg Thieme Verlag, Stuttgart, Germany.
Schaltenbrand, G., Wahren, W., 1977. Atlas for Stereotaxy of the Human Brain. Georg Thieme Verlag, Stuttgart, Germany.
St-Jean, P., Sadikot, A.F., Collins, L., Clonda, D., Kasrai, R., Evans, A.C., Peters, T.M., 1998. Automated atlas integration and interactive three-dimensional visualization tools for planning and guidance in functional neurosurgery. IEEE Transactions on Medical Imaging 17 (5), 672–680.
Starr, P.A., Vitek, J.L., DeLong, M., Bakay, R.A., 1999. Magnetic resonance imaging-based stereotactic localization of the globus pallidus and subthalamic nucleus. Neurosurgery 44 (2), 303–314.
Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Georg Thieme Verlag, Stuttgart, Germany.
Thirion, J.-P., 1996. Non-rigid matching using demons. In: Computer Vision and Pattern Recognition, CVPR'96, San Francisco, CA, USA, June. Electronic version: <http://www.inria.fr/RRRT/RR-2547.html>.
Thirion, J.-P., 1998. Image matching as a diffusion process: an analogy with Maxwell's demons. Medical Image Analysis 2 (3), 243–260.
Vidailhet, M., Vercueil, L., Houeto, J.L., Krystowiak, P., Lagrange, C., Yelnik, J., Bardinet, E., Benabid, A.L., Navarro, S., Dormont, D., Grand, S., Blond, S., Ardouin, C., Pillon, B., Dujardin, K., Hahn-Barma, V., Agid, Y., Destée, A., Pollak, P.and the French SPIDY Study Group, 2007. Bilateral, pallidal, deep-brain stimulation in primary generalised dystonia: a prospective 3 year follow. Lancet Neurology 6 (3), 223–229.
Warfield, S.K., Zou, K.H., Wells, W.M., 2004. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. IEEE Transactions in Medical Imaging 23 (7), 903–921.
Xu, M., Nowinski, W.L., 2001. Talairach–Tournoux brain atlas registration using a metalforming principle-based finite element method. Medical Image Analysis 5 (4), 271–279.
Yelnik, J., Bardinet, E., Dormont, D., Malandain, G., Ourselin, S., Tandé, D., Karachi, C., Ayache, N., Cornu, P., Agid, Y., 2007. A three dimensional, histological and deformable atlas of the human basal ganglia. I. Atlas construction based on immunohistochemical and MRI data. NeuroImage 34 (2), 618–638.