Poster presented at the 11th Annual Meeting of the Organization for Human Brain Mapping, June 12-16th, 2005 in Toronto, Canada.

Inter-rater Reproducibility of 3D Cortical and Sub-cortical Landmark Points

Kristi Boesen[1], Stephen Frey[3], Jill Huang[1], Jurgen Germann[3], Joshua Stern[1], D. Louis Collins[3], Alan C. Evans[3], David A. Rottenberg[1,2]

[1]University of Minnesota Department of Neurology
[2]University of Minnesota Department of Radiology
[3]McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University

www.neurovia.umn.edu/INC

www.nimh.nih.gov/neuroinformatics

INTRODUCTION

Accurate, reproducible landmarking of the cortical surface is difficult due to variability in individual patterns of sulcation and gyrification, and no previous study has generated cortical landmarks with sufficient inter-rater reproducibility to be useful for the quantitative evaluation of intersubject registration algorithms.

Two recent attempts at landmarking [1,2] employed reconstructed cortical surfaces with landmarks based on visible sulcal intersections or visually extrapolated intersections assumed to lie on the adjacent cortical surface. In the first of these attempts 12 x 2 landmarks (12 in each hemisphere) were derived from the intersections of 18 x 2 homologous sulci and identified according to a written protocol by two independent raters in 20 subject volumes. Reproducibility of these 12 x 2 landmarks was then assessed in six subject volumes: landmarks were identified repeatedly by one of the former raters and once by a third independent rater. For these two raters, the Euclidean distance between individual and mean landmark positions for 12 x 2 landmarks averaged 6.0 mm; intra-rater variability averaged 3.7 mm [1]. In the second of these attempts, the same 12 x 2 landmarks from [1] were identified using a revised written protocol in 58 subject volumes [2]. Repeatability based on replicate measurements by a single rater ranged from 0.25 to 1.55 mm. In both studies the distribution of landmarks included few landmarks on the parietal and occipital cortices, and none on the temporal cortex.

In another study [3] 128 x 2 landmarks were identified by either of two raters in seven subject volumes; landmarks were defined based on ten anatomic reference planes with 118 x 2 landmarks located specifically on the reference planes. Intra-rater reliability of these points involving a single repeat labeling of each landmark on a single reference brain yielded a RMS error of 1.6 mm. No inter-rater reliability results were reported.

METHODS

Four raters at UM and MNI independently identified 13 x 2 cortical and 10 sub-cortical point landmarks on 37 high-quality 1.5T T1-weighted MRI volumes (voxel dimensions 1 x 1 x 1 mm) based on a detailed written protocol. Prior to landmarking MRI volumes were corrected for intensity non-uniformity, intensity normalized, and transformed into unscaled Talairach space [4]. Using REGISTER visualization software (MNI) landmarks were identified by 3D navigation within the MRI volumes and their (x, y, z) coordinates were recorded for subsequent statistical analysis.

Landmarks were chosen for their usefulness in determining Goodness of Warp (GoW) metrics for subsequent evaluation of intersubject registration algorithms. Cortical landmark points (13 x 2) were widely distributed among surfaces and sulci in the frontal and temporal lobes including 3 x 2 landmark points on the temporal-insular transition in the coronal plane. Few cortical landmark points were identified in the occipital lobe, and none were identified in the parietal lobe as structural homology is difficult to establish in these regions. Sub-cortical landmark points (10) were distributed along the AC-PC line/midsagittal plane and at the anterior and posterior regions of the lateral ventricles in each hemisphere (Table 1).

To calculate inter-rater reproducibility we computed the distance from each of the four raters' points to the centroid of those four points for each landmark/volume combination, yielding 5,328 distances. Protocols for 6 x 2 landmarks were revised, and the resulting 1,776 points were re-labeled. To minimize the effect of possible careless errors, an "outlier index" was computed for each landmark/volume combination: the three distances between a given rater's point and those of the other three raters were averaged and subtracted from the mean of the three distances between the other three raters. Possible outliers were identified as those having an outlier index > 2 mm, and raters were asked to re-examine their own outlier points without being told how to adjust them. Of the 433 possible outliers identified, 233 were adjusted leading to a mean improvement of 10.5 mm with respect to the centroid of the other three raters.

METHODS

Cortical Points			
5,6	Inferior frontal s.	21,22	Superior frontal s. caudal
9,10	Temporal lobe anterior	23,24	Superior frontal s. rostral
11,12	Calcarine fissure anterior	25,26	Olfactory s. rostral
13,14	Occipital pole posterior	31,32	Sylvian fissure
15,16	Frontal pole anterior	33,34	Superior temporal s.
17,18	Central s. dorsal	35,36	Collateral s.
19,20	Central s. ventral		
Sub-Cortical Points			
1	AC	27	CC anterior
2	PC	28	CC posterior
3,4	Lateral ventricle anterior	29	CC ventral extent of splenium
7,8	Lateral ventricle posterior	30	Thalamus center

Table 1. Thirty-six cortical and sub-cortical landmark points. Paired numbers represent landmark points labeled in both hemispheres, left and right, respectively. Abbreviations: s., sulcus; AC, anterior commissure; PC, posterior commissure; CC, corpus callosum.

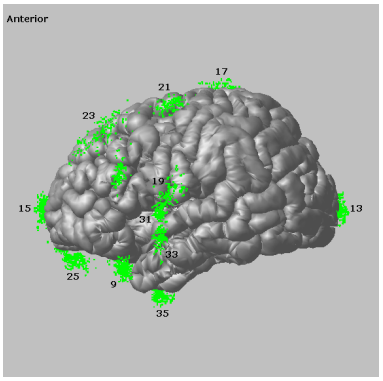

Figure 1. Lateral view of cortical landmark points in the left hemisphere (surface shown is for visual reference). Point clouds in green represent Talairach coordinates of points labeled by four raters over 37 volumes for each landmark. See Table 1 for landmark labels.

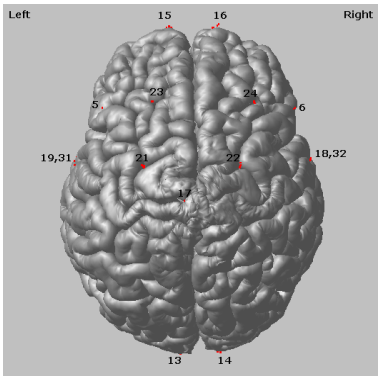

Figure 2. Dorsal view of cortical landmark points in both hemispheres of a single subject volume. Point clouds in red represent Talairach coordinates of points labeled by four raters for this volume, with a small point cloud indicating overlap between raters. See Table 1 for landmark labels.

RESULTS

Preliminary reproducibility statistics for the 5,328 distances yielded a median distance to centroid for cortical and non-cortical landmark points, respectively, of 1.24 and 0.68 mm. After identifying possible outliers, defined as those having an outlier index > 2 mm, the median distance to centroid for cortical and non-cortical landmark points, respectively, was 1.06 and 0.63 mm. Reproducibility statistics for the final set of 5,328 distances are listed in Tables 2 and 3.

RESULTS

Cortical Points	L Median	R Median	L Mean	R Mean
Inferior frontal s.	0.78	0.90	1.44	1.19
Occipital pole posterior	0.79	0.78	0.87	0.84
Frontal pole anterior	1.02	1.30	1.19	1.42
Temporal lobe anterior	1.06	1.04	1.30	1.23
Calcarine fissure anterior	1.19	1.19	1.34	1.25
Central s. dorsal	0.67	0.67	0.76	0.75
Central s. ventral	1.28	1.67	1.50	2.04
Olfactory s. rostral	0.83	0.86	1.02	0.96
Sylvian fissure	1.21	1.27	1.26	1.35
Superior temporal s.	1.24	1.34	1.46	1.64
Collateral s.	1.28	1.93	1.80	2.57
Superior frontal s. caudal	1.47	1.32	2.87	1.73
Superior frontal s. rostral	4.75	3.13	6.19	4.49

Table 2. Reproducibility statistics (mm) describing median and distance to centroid and mean distance to centroid for cortical landmark points over 37 volumes by four raters. Abbreviations: L, left; R, right; s., sulcus.

Subcortical Points	Median	Mean
AC point	0.59	0.61
PC point	0.61	0.67
Lateral ventricle anterior L	0.57	0.59
Lateral ventricle anterior R	0.59	0.61
Lateral ventricle posterior L	0.70	1.81
Lateral ventricle posterior R	0.65	1.70
CC anterior	0.79	0.80
CC posterior	0.81	0.82
CC ventral extent of splenium	0.71	0.73
Thalamus center	0.70	0.71

Table 3. Reproducibility statistics (mm) describing median distance to centroid and mean distance to centroid for sub-cortical landmark points over 37 volumes by four raters. Abbreviations: AC, anterior commissure; PC, posterior commissure; L, left; R, right; CC, corpus callosum.

DISCUSSION

Our results suggest that a large number of distributed cortical and sub-cortical landmarks are highly reproducible across raters, sites, and subjects (Tables 2 and 3). In addition, point clouds consisting of label points from 37 subjects by four raters for each landmark point provide visual evidence for the goodness of the affine registration performed on the subject volumes during pre-processing (Figs. 1 and 2).

As previously stated, variability in individual sulcal and gyral patterns may pose limits on reproducibility across subjects; the landmark points labeling the rostral extent of the left and right superior frontal sulci had median and mean distances to centroid greater than the outlier-defining distance of 2 mm (Table 2) and serve as examples of this variability across subjects.

Over the entire group of landmark/volume combinations and raters in this study, reproducibility statistics yielded overall median and mean distances to centroid for cortical and sub-cortical landmark points of 1.06 and 0.63 mm, respectively, and we conclude that this set of landmarks points can serve as well-characterized points of reference for future developmental, neuroanatomical, and disease-related studies as well as goodness-of-warp metrics.

REFERENCES

1. Maudgil DD, *et al.* J. Anat 193: 559-571, 1998.
2. Free SL, *et al.* NeuroImage 13:801-813, 2001.
3. Grachev ID, *et al.* NeuroImage 9: 250-268, 1999.
4. Collins DL, *et al.* JCAT 18:192-205, 1994.

This work is supported in part by NIH Grant P20 EB002013 (Human Brain Project).

The authors would like to thank Dr. Michael Petrides for lending his expertise and insight in discussions regarding neuroanatomy of the frontal cortex.

Comments or questions may be directed to Kristi Boesen via e-mail at kristi@neurovia.umn.edu