Supplementary sections to accompany:
"Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration"
Arno Klein, et al., NeuroImage (2009) [doi:10.1016/j.neuroimage.2008.12.037]
http://www.mindboggle.info/papers

1 Supplementary material: Initial run

A few algorithms resulted in consistently low accuracies or occasional failures in their registrations. We did our best to rectify these problems. For example, ROMEO's initial run produced inconsistent results, so Hellier provided intensity correction code which we applied to all of the data prior to registering with ROMEO. Of the 2,168 registrations, there were 19 failures for AIR, accompanied by the message: "Registration terminated due to a Hessian matrix that was not positive definite" (even with the "-q" option). We were able to correct all of these cases by skipping the "alignlinear" step (and relying on the preliminary linear alignment with FLIRT). SyN had seven failures, which we corrected by first running Avant's TranslateRegistration program (this program had no effect on other registrations).

SPM's DARTEL Toolbox resulted in highly variable results for the LPBA40 set (and low results obtained with one of the MGH10 images), most likely because of the inconsistent way that an older version of the code dealt with zeros in the images. DARTEL estimates its spatial transformation by registering gray and white matter maps produced by the segmentation step. Regions in these images, which had been set to zero in the skull-stripped data, sometimes contained information from the tissue probability maps used by the segmentation. We downloaded an updated version of DARTEL that corrects for this and ran it again (in both a pairwise as well as average template manner).

In addition to addressing variable results, we had to set reasonable time constraints. SyN's default parameters were found to be too computationally intensive, so Avants recommended a different number of iterations to keep computation under an hour per registration. Likewise, Rueckert provided a parameter file with reduced control point spacing to reduce IRTK's computation time. IRTK, SICLE, and SPM's Unified Segmentation and DARTEL Toolbox were run by their authors on one registration pair to ensure that the first author's implementation and execution of their software was correct. [1]

[1] Likewise for PASHA and HAMMER, but these were excluded from the study (see Discussion).

2 Supplementary material: The trivial case: self-registration results

For the trivial case, where each brain was registered to itself, all of the methods performed nearly perfectly as measured by volume similarity, with the exception of SPM's DARTEL Toolbox for the LPBA40 set. According to target volume overlap, the methods that gave less than perfect results were: ANIMAL and ROMEO (for all four label sets), "SPM2-type" Normalization (IBSR18, CUMC12, and MGH10), and ROMEO, SICLE, and SPM's DARTEL (LPBA40). The few cases of low values were obtained with SPM's DARTEL, and only for the LPBA40 set according to overlap and distance measures.

The results were almost identical for the target, union, and mean volume and surface overlap measures. Not surprisingly, the deviations from perfect registration in the trivial case are revisited in the non-trivial, inter-brain registrations results.

3 Supplementary material: More indifference-zone rankings

See Figures 8, 9, and 10 for indifference-zone ranking of the registration methods for the IBSR18, CUMC12, and MGH10 data, respectively. These rankings are based on target overlap results, and correspond to the indifference-zone ranking for the LPBA40 data in Figure 7.

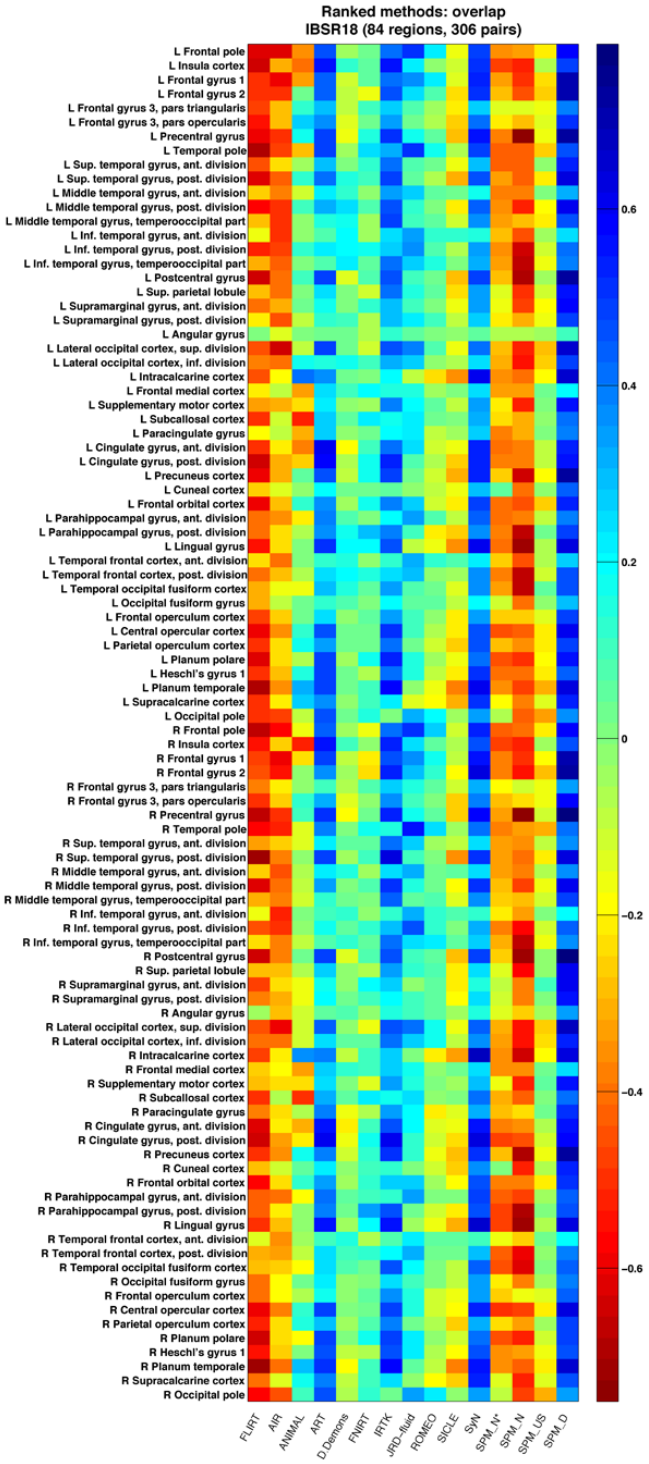

Fig. 8. Indifference-zone ranking of the registration methods: IBSR18 overlaps. This matrix was constructed as in Figure 7 for target overlap rankings averaged across 306 registration pairs using the 84 regions of the IBSR18 dataset (union and mean overlap results are almost identical). Blue indicates higher accuracy; the colors (and color range) are not comparable to those of the other label sets (Figures 7, 9, and 10). (SPM_N*="SPM2-type" Normalize, SPM_N=Normalize, SPM_US=Unified Segmentation, SPM_D=DARTEL pairwise)

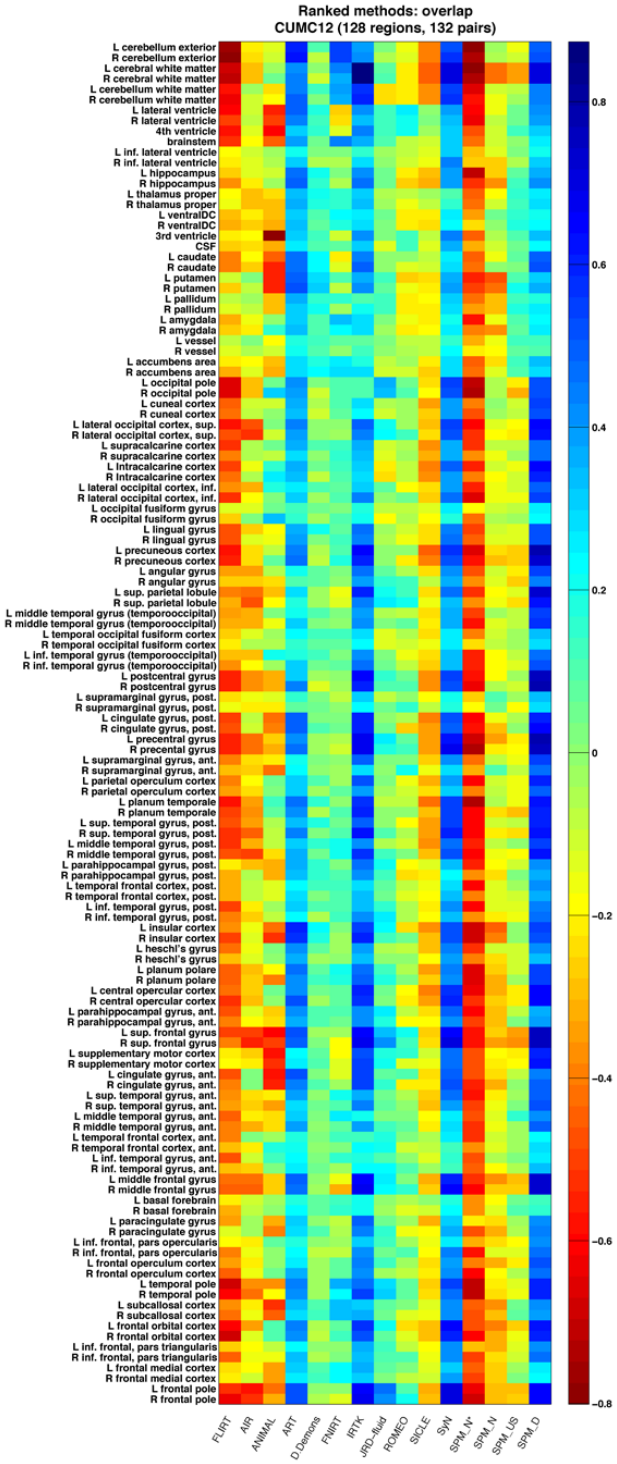

Fig. 9. Indifference-zone ranking of the registration methods: CUMC12 overlaps. This matrix was constructed as in Figure 7 for target overlap rankings averaged across 132 registration pairs using the 128 regions of the CUMC12 dataset (union and mean overlap results are almost identical). Blue indicates higher accuracy; the colors (and color range) are not comparable to those of the other label sets (Figures 7, 8, and 10). (SPM_N*="SPM2-type" Normalize, SPM_N=Normalize, SPM_US=Unified Segmentation, SPM_D=DARTEL pairwise)

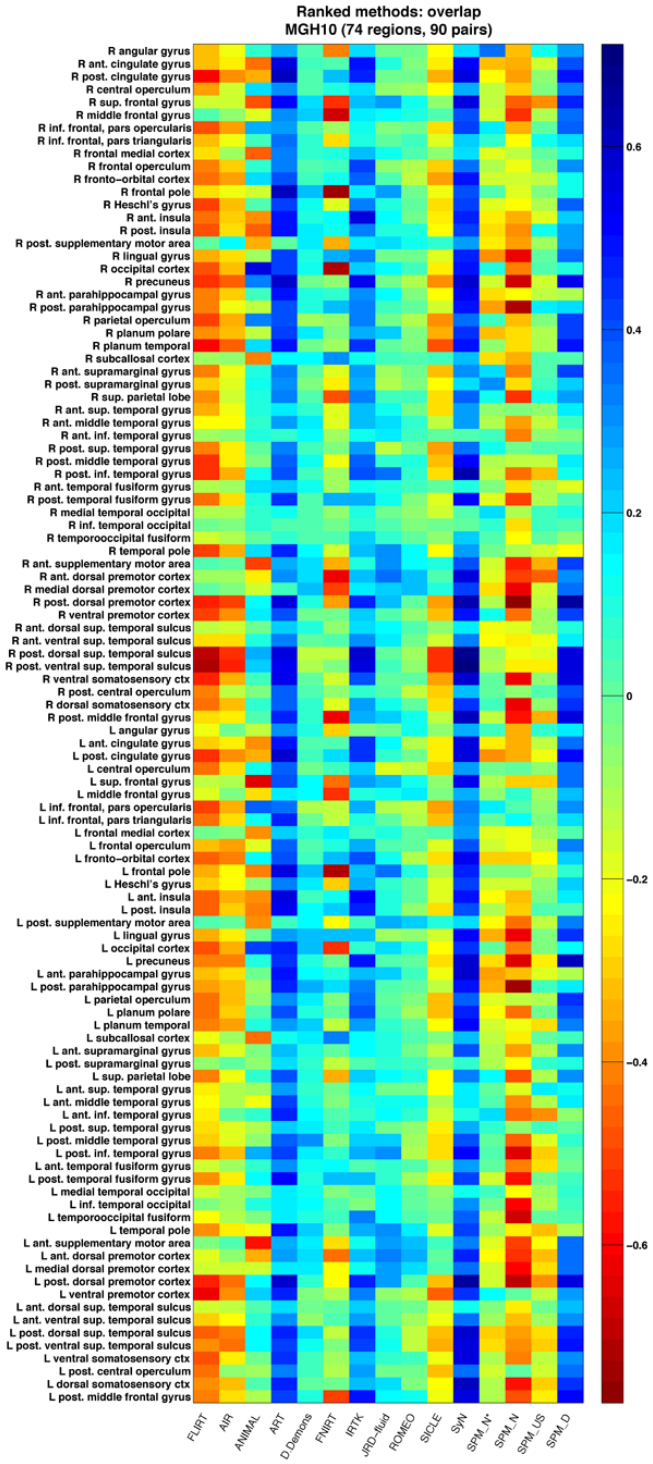

Fig. 10. Indifference-zone ranking of the registration methods: MGH10 overlaps. This matrix was constructed as in Figure 7 for target overlap rankings averaged across 90 registration pairs using the 74 regions of the MGH10 dataset (union and mean overlap results are almost identical). Blue indicates higher accuracy; the colors (and color range) are not comparable to those of the other label sets (Figures 7, 8, and 9). (SPM_N*="SPM2-type" Normalize, SPM_N=Normalize, SPM_US=Unified Segmentation, SPM_D=DARTEL pairwise)

4 Supplementary material: Volume similarity results

The volume similarity measures, averaged across all of the regions in each label set, are shown in Figure 11. Volume similarity is only an indirect measure of registration accuracy, however it can be used to expose gross discrepancies in volume, if not shape. For each label set, all of the methods resulted in very similar median values. The LPBA40 set and SPM DARTEL exhibited the largest number of outliers.

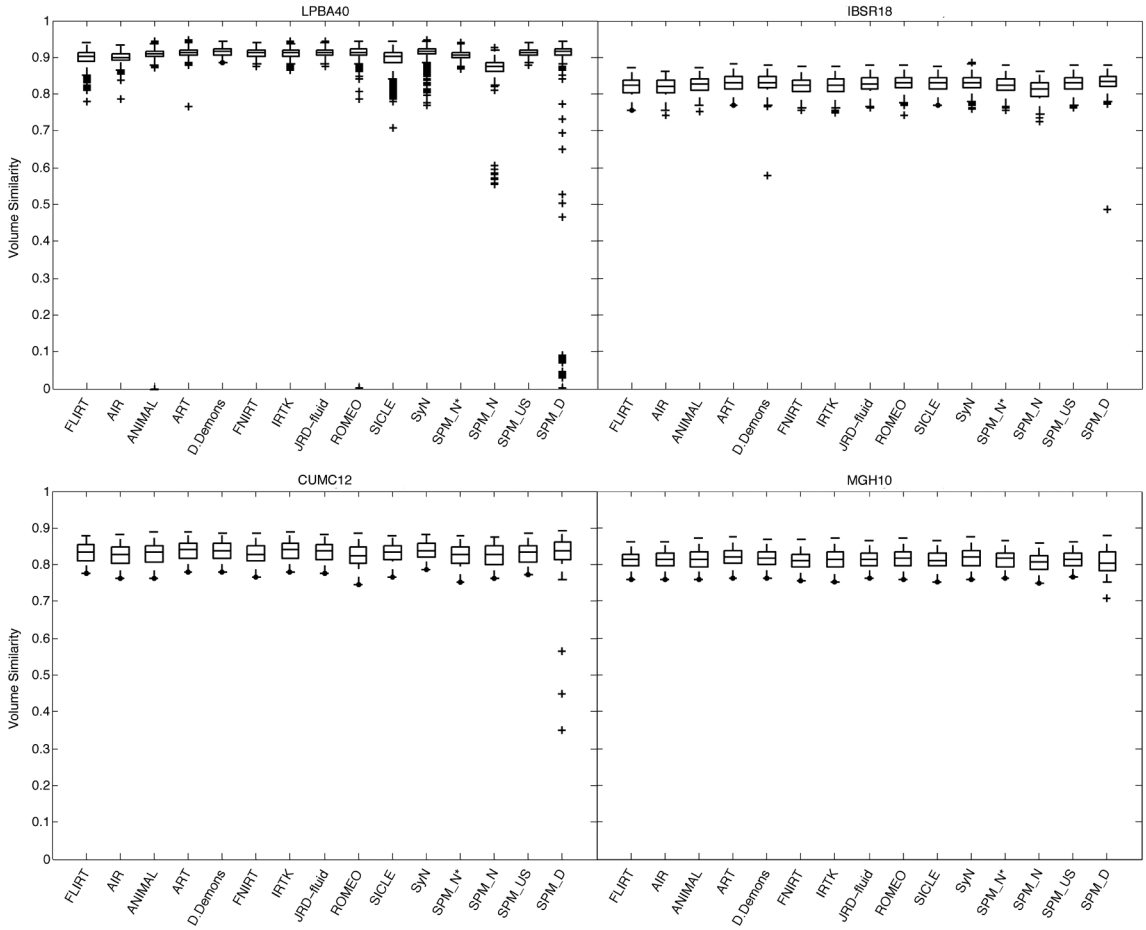

Fig. 11. Volume similarity by registration method. These box and whisker plots (constructed as in Figure 5) show the volume similarity measures between deformed source and target labels, averaged first across all of the regions in each label set (LPBA40, IBSR18, CUMC12, and MGH10) then across brain pairs, with highest similarity at the top. (SPM_N*="SPM2-type" Normalize, SPM_N=Normalize, SPM_US=Unified Segmentation, SPM_D=DARTEL pairwise)

5 Supplementary material: Distance results

The box and whisker plot in Figure 12 shows the distance errors for the LPBA40 label set. They exhibit a roughly similar pattern across the methods as the overlap values in Figure 5, although regional differences may be seen across Figures 13 and 7.

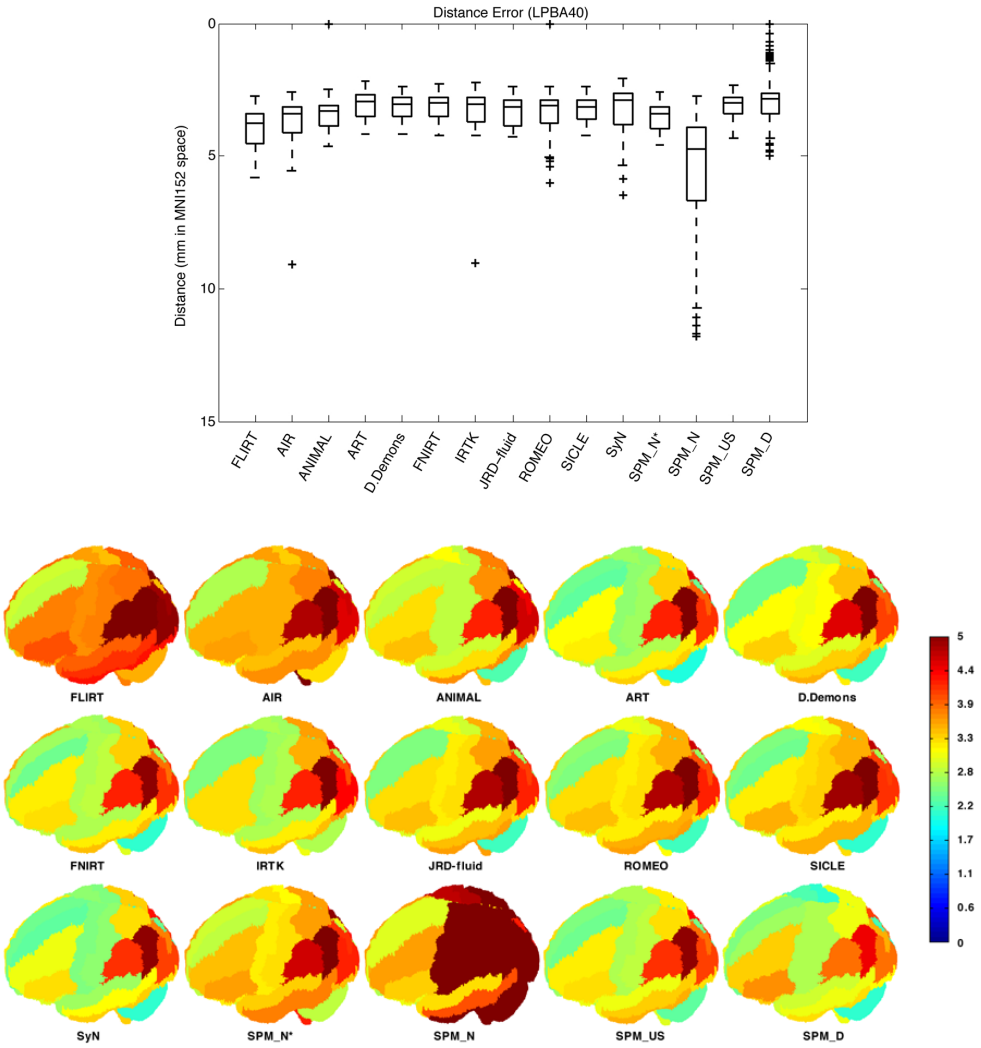

Fig. 12. Distance error by registration method. The box and whisker plot was constructed as in Figure 5 except that the measure is distance error between deformed source and target label boundaries, averaged first across all of the regions in the LPBA40 label set then across brain pairs, with lowest errors toward the top. The brain images (constructed as in Figure 6) show the mean distance error per region as a color (blue indicates higher accuracy). (SPM_N*="SPM2-type" normalization, SPM_N=SPM's Normalize, SPM_US=Unified Segmentation, SPM_D=DARTEL pairwise)

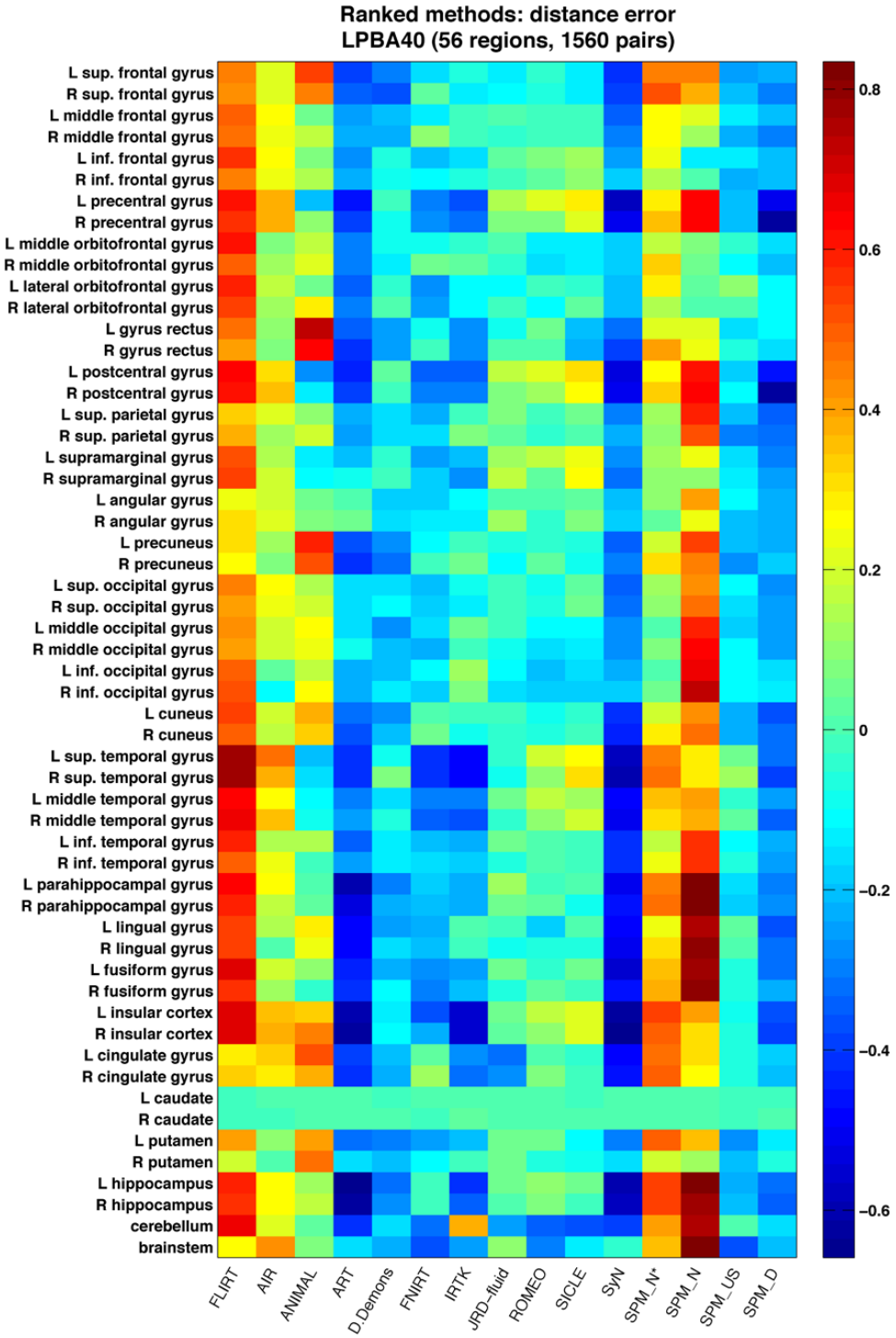

Fig. 13. Indifference-zone ranking of the registration methods: LPBA40 distance errors. This matrix was constructed as in Figure 7, except for distance error rankings rather than overlap rankings. Blue indicates lower error and higher accuracy. (SPM_N*="SPM2-type" normalization, SPM_N=SPM's Normalize, SPM_US=Unified Segmentation, SPM_D=DARTEL pairwise)

6 Supplementary material: Dependence test

To determine the degree of correlation across source-target registration pairs, we selected and organized pairs from the LPBA40 brains (e.g., a→b) into two columns (a→b, c→d), where rows were independent of one another (no brain appeared in multiple rows), but the two columns had one of four dependency relationships to one another, as follows:

(1) No dependence; each brain was used only once (10 rows):
    a→b, c→d
(2) source and target dependence; both brains were used twice (20 rows):
    a→b, b→a
(3) source dependence; the source is used twice (13 rows):
    a→b, a→c
(4) target dependence; the target is used twice (13 rows):
    a→b, c→b

We then replaced each registration pair with the target overlap calculated for that pair for each registration method, and computed the correlation between the left and right columns for each dependency condition. We repeated this procedure 1,000 times, using a new set of pairs satisfying each condition, and averaged the results. The values for the no dependence condition are close to zero as we would expect for independent pairs. All of the other conditions, however, result in high correlations for most of the methods. One interesting detail is that there is an asymmetry between conditions 3 and 4. For example, for some methods, even when pairs that share a source brain are correlated, pairs that share a target brain are not necessarily correlated.

7 Supplementary material: Software setup and commands

Most of the programs offer a range of flexible options for their similarity metric, regularization method, etc. Therefore the following commands each represent but a single implementation of its underlying algorithm. The nonlinear deformation programs are listed in alphabetical order after the linear registration program FLIRT.

Python and Matlab programs were used to call all of the commands below, and were run on an OSX system (Mac Pro 2-Quad-Core (8-processor) Intel Xeon, 3GHz, 6GB RAM) with a 10.4 operating system, except for ROMEO (10.5 operating system, same hardware), ART (Dell PowerEdge 6600 Enterprise server with four 2.8GHz Intel Xeon processors and 28GB of RAM running Redhat linux), and JRD-fluid (run on LONI's servers: SUN Microsystem workstations with a dual 64-bit AMD Opteron 2.4 GHz processor running Solaris). IRTK was run on some of the LPBA40 brains on servers at Imperial College, and the results were confirmed to be identical to those obtained by the first author.

FLIRT linear and rigid alignment commands were run before any of the other commands, and the resulting linearly transformed brain images were resliced using trilinear interpolation. For all algorithms, manual labels were resliced using nearest-neighbor interpolation.

FLIRT (FSL v4.0)
*Linearly align **to** MNI space (9-parameter):*
flirt -in ⟨source⟩ -ref ⟨MNI152⟩ -omat ⟨output transform⟩.mat -out ⟨output brain⟩ -bins 256 -cost corratio -searchrx -90 90 -searchry -90 90 -searchrz -90 90 -dof 9 -interp trilinear
where ⟨MNI152⟩ is the nonlinear MNI152 template used by FSL: MNI152_T1_1mm_brain

*Rigidly align **in** MNI space (6-parameter):*
flirt -in ⟨above output brain⟩ -ref ⟨target⟩ -omat ⟨output transform⟩.mat -out ⟨starting point for all other algorithms⟩ -bins 256 -cost corratio -searchrx -30 30 -searchry -30 30 -searchrz -30 30 -dof 6 -interp trilinear

Reslice using 6-parameter output transform:
flirt -in ⟨labeled source⟩ -ref ⟨labeled target⟩ -applyxfm -init ⟨transform⟩.mat -out ⟨output labels⟩ -paddingsize 0.0 -interp nearestneighbour

AIR v5.25
Align (12-parameter affine):
alignlinear ⟨target⟩.img ⟨source⟩.img ⟨output affine transform⟩.air -m 12 -t1 1 -t2 1 -q

Warp (2nd-order, 3rd-order, 4th-order and then 5th-order 168-parameter):
align_warp ⟨target⟩.img ⟨source⟩.img ⟨output transform⟩.warp -m 2 5 -f ⟨above affine transform⟩ -t1 1 -t2 1 -q

Reslice:
reslice_warp ⟨align_warp transform⟩.warp ⟨output labels⟩.img -a ⟨labeled source⟩.img -n 0 -o

ANIMAL (MNI AutoReg v0.98k)
A Perl script performs non-linear fitting. The original ANIMAL parameters were optimized Robbins (**?**). The resulting Perl script was modified again by Janke and Lepage. The script calls minctracc and the parameters and hierarchical steps are as follows:

'-nonlinear', 'corrcoeff', '-weight',1, '-stiffness',1, '-similarity',0.3, '-sub_lattice',6

'step':	32	16	12	8	6	4
'blur_fwhm':	16	8	6	4	3	2
'iterations':	20	20	20	20	20	10

Warp: ⟨Perl script⟩.pl -clobber -normalize ⟨source⟩.mnc ⟨target⟩.mnc ⟨output transform⟩.xfm

Reslice: mincresample -transform ⟨transform⟩.xfm -like ⟨labeled target⟩.mnc ⟨source⟩.mnc ⟨output labels⟩.mnc -nearest_neighbour -keep_real_range -short

Convert to Analyze format: mnc2nii -short ⟨above output labels⟩.mnc ⟨output labels⟩.nii

ART
Warp: 3dwarper -R -trg ⟨target⟩.img -obj ⟨source⟩ -u ⟨output transform⟩.wrp -o ⟨output brain⟩.img -A -sd 8.0

Reslice: applywarp3d -nn -w ⟨transform⟩.wrp ⟨labeled source⟩

Diffeomorphic Demons:

Warp: DemonsRegistration -f ⟨target⟩.hdr -m ⟨source⟩.hdr -0 ⟨output transform⟩.hdr -e -s2 -i30x20x10

Reslice: ResampleImage -i ⟨labeled source⟩.hdr -f ⟨transform⟩.hdr -o ⟨output labels⟩.img

FNIRT
Warp: fnirt --config=schedule_01.cnf --ref=⟨target⟩ --in=⟨source⟩ --cout ⟨output transform⟩

Reslice: applywarp -i ⟨labeled source⟩ -r ⟨target⟩ -w ⟨transform⟩ -o ⟨output labels⟩ --interp=nn

IRTK
In the current study no regularization was used.
The parameter file called for a 2.5 mm minimum control point spacing, and the non-rigid registration parameters were:
Lambda1, 2, and 3 = 0
Control point spacing in X, Y, and Z = 20

Rigidly align: rreg ⟨target⟩.hdr ⟨source⟩.hdr -dofout ⟨output transform⟩_rreg.dof

Affinely align: areg ⟨target⟩.hdr ⟨source⟩.hdr -dofin ⟨rreg transform⟩_rreg.dof -dofout ⟨output transform⟩_areg.dof

Warp: nreg ⟨target⟩.hdr ⟨source⟩.hdr -dofin ⟨areg transform⟩_areg.dof -dofout ⟨output transform⟩_nreg.dof -parin ITK_parameters_2.5mm.txt

Reslice: transformation ⟨labeled source⟩.hdr ⟨output labels⟩.hdr -dofin ⟨nreg transform⟩_nreg.dof -target ⟨target⟩.hdr

JRD-Fluid:
Warp: mix_fluidmap_nohassle_column.out ⟨source⟩.img ⟨dimensions⟩ ⟨target⟩.img ⟨dimensions⟩ ⟨output directory⟩ source_le 0.95 0 ⟨parameter⟩,

where *parameter* is set to 4 for LONI LPBA40 images (181x217x181 voxels) and set to 20 for all other images registered to the nonlinear MNI152 template.

Reslice: resample_8bit_image_NN.out ⟨labeled source⟩.img ⟨above output directory⟩/trn_DXP_ftle ⟨above output directory⟩/trn_DYP_ftle ⟨above output directory⟩/trn_DZP_ftle ⟨dimensions⟩ ⟨output labels⟩

ROMEO
Warp and reslice: Romeo.py -t ⟨target⟩ -s ⟨source⟩ -l ⟨labeled source⟩ -o ⟨output brain⟩ -d ⟨output labels⟩

SICLE
Preparation: lereg2.exe -gf global.param -p input_files/⟨source-target pair parameter file⟩.in | tee ⟨source-target pair output file⟩.out

Warp: rpp2.exe ⟨source-target pair parameter file⟩.in -def -jac -gf global.param

Reslice: deform3d.exe -coeff ⟨source-target pair output stem⟩_res10000_iter00020.coeffs -t ⟨labeled source⟩.hdr -o ⟨output labels⟩.hdr -interp N

SPM5's "SPM2-type" Normalization
Warp (in Matlab): spm_normalise("⟨target⟩.img", "⟨source⟩.img", "⟨output transform⟩.mat");

Reslice (in Matlab): spm_write_sn("⟨labeled source⟩.img", "⟨output labels⟩.mat", struct("interp",0));

SPM5's Normalize, Unified Segmentation, and DARTEL Toolbox
Matlab scripts were used to compose the transforms for these methods, and may be obtained from `http://www.mindboggle.info/papers/`.

SyN
Linearly align (optional): TranslationRegistration ⟨target⟩.hdr ⟨source⟩.hdr ⟨output linear transform⟩

Warp: perl arnoMVSN.pl ⟨target⟩.hdr ⟨source⟩.hdr ⟨above linear transform⟩ " -c 5 -n3 -i30x99x11 -l 0.5 -s 2 -a 0.05 " ⟨SyN directory⟩ 0 0 0 ⟨user name⟩ 1

Reslice: WarpImageBackward ⟨labeled source⟩.hdr ⟨transform filestem⟩warp ⟨output labels⟩.hdr donearestneighbor

We corroborated results obtained with the above commands for the newer, publicly released ANTS software (`http://www.picsl.upenn.edu/ANTS/`) using the following commands:

Warp: ANTS 3 -m PR[⟨target⟩.nii, ⟨source⟩.nii, 1, 2] -o ⟨output transform⟩.nii -r Gauss[2,0] -t SyN[0.5] -i 30x99x11 –use-Histogram-Matching

Reslice: WarpImageMultiTransform 3 ⟨labeled source⟩.nii ⟨output labels⟩.nii -R ⟨target⟩.nii ⟨transform⟩Warp.nii ⟨transform⟩Affine.txt –use-NN

8 Supplementary material: Algorithm descriptions

8.1 FLIRT: FMRIB's Linear Image Registration Tool

FLIRT was developed by Jenkinson (**?**) at the FMRIB Centre at the University of Oxford, UK. FLIRT can be downloaded with the FMRIB Software Library (FSL) at `http://www.fmrib.ox.ac.uk/fsl/`.

FLIRT is an automated linear (affine) registration tool based around a multi-start, multi-resolution global optimisation method. It can be used for inter- and intra-modal registration with 2-D or 3-D images. In addition, it can be run with a number of different transformation models (degrees of freedom) and it implements a general cost function weighting scheme for all cost functions.

8.2 AIR: Automated Image Registration

AIR was developed by Woods (**?**) at the David Geffen School of Medicine at UCLA, CA. The C source code may be downloaded at `http://air.bmap.ucla.edu`.

AIR aligns a pair of images by minimizing the mean squared difference between the image designated for resampling and the image to which it is being registered. First and second derivatives of the cost function are computed and used to iteratively adjust the model parameters. In addition to an optional intensity scaling parameter, the algorithm adjusts the elements of polynomials of the desired order. Registration begins with a first order polynomial (i.e., an affine transformation) and the order is incremented sequentially until the order specified by the user is reached. Polynomial orders up to twelfth order are implemented. At each order, registration begins with sparse sampling of the data and proceeds to denser sampling, per user specifications, with subsequent iterations. Criteria for advancing to denser sampling or to higher order polynomials include the magnitude of the predicted improvement in the cost function, the number of iterations without improvement in the cost function, and the total number of iterations. In some instances, full Newton-based minimization cannot proceed because the Hessian matrix of second derivatives is not positive definite. The likelihood of a non-positive definite Hessian matrix can optionally be reduced by omitting certain contributions to the second derivatives using a strategy similar to that employed by the Levenberg-Marquardt algorithm. The optimized polynomial transformation is stored, and separate programs allow either of the two original images to be resampled into the space defined by the other image using a variety of interpolation algorithms (nearest neighbor interpolation, trilinear interpolation, sinc interpolation, and

hybrids of these methods). Diagnostics are available to verify that no regions with non-positive Jacobians are present. Polynomial transformations can be mathematically combined with any number of affine linear transformations, allowing data to be resampled directly from or to any space that has an affine relationship to the original images.

(degrees of freedom (dof) = 168)

8.3 ANIMAL: Automatic Nonlinear Image Matching and Anatomical Labeling

ANIMAL was developed by Collins et al. (**????**) of the Montreal Neurological Institute, Canada. The original ANIMAL is available for download at `http://www.bic.mni.mcgill.ca/users/louis/MNI_ANIMAL_home/readme/`. Lepage provided a Perl script (see below) for implementing the multi-resolution strategy.

ANIMAL is based on multi-scale, 3-D cross-correlation. Spatial registration is completed automatically as a two step process. The first accounts for the linear part of the transformation by using correlation between Gaussian-blurred features extracted from both volumes. In the second step, ANIMAL estimates the 3-D deformation field required to account for this variability. The deformation field is built by sequentially stepping through the target volume in a 3-D grid pattern. At each grid-node i, the deformation vector required to achieve local registration between the two volumes is found by optimization of three translational parameters (txi, tyi, tzi) that maximize the objective function evaluated only in the neighborhood region surrounding the node. The algorithm is applied iteratively in a multi-scale hierarchy, so that image blurring and grid size are reduced after each iteration, thus refining the fit. The multi-scale approach also makes the procedure very robust and ensures that the algorithm converges to the global minimum.

A Perl script (nlfit) implements the multi-resolution fitting strategy to map brains into stereotaxic space at the Montreal Neurological Institute. At the heart of this procedure is minctracc, the program that automatically finds the best non-linear transformation to map one volumetric data set (stored in MINC format) to another. The program uses optimization over a user selectable number of parameters to identify the best transformation mapping voxel values of the first data set into the second.

(dof $\leq 3 \times \simeq$23,000 nodes (1.5M brain voxels / 4mm each direction) = 69,000)

8.4 ART: Automated Registration Tool

ART was developed by Ardekani et al. (**?**) at the Nathan Kline Institute, NY. The executables may can be downloaded at `http://www.nitrc.org/projects/art/`. Ardekani revised the registration program to allow one to specify an output file and revised the resampling code to enable nearest-neighbor interpolation.

ART uses local normalized cross-correlation between the source and target images as its similarity measure. It determines a displacement vector field defined for each grid point on the target image using a non-parametric free-form multi-resolution approach. The displacement vector field obtained at each resolution level is regularized by median and low-pass filtering.

Prior to non-linear registration, there are options for the program to determine a linear rigid-body (6-parameter) registration between the target and subject images followed by a linear affine (12-parameter) registration (**?**).

(dof $\simeq$ 3 $\times$#brain voxels $\simeq$ 7million)

8.5 Demons: Diffeomorphic Demons

Diffeomorphic Demons was developed by Vercauteren et al. (**?**) at Mauna Kea Technologies and at INRIA Sophia Antipolis, France; it is implemented as part of the finite difference solver framework within the Insight Toolkit (ITK) and can be downloaded at `http://hdl.handle.net/1926/510`. A graphical user interface is provided as part of MedINRIA (**?**). Vercauteren provided a version of the resampling code with an option for nearest-neighbor interpolation that doesn't require origin information.

This non-parametric algorithm generalizes Thirion's Demons algorithm (**?**) to produce a diffeomorphic spatial transformation (**?**). This method alternates between the computation of warping forces inspired from optical flow theory and the regularization of these forces by a simple Gaussian smoothing.

The Demons algorithm may be characterized as the optimization of a global energy function (Cachier et al., 2003) where correspondences act as a hidden variable in the registration process. The regularization criterion is then considered as a prior on the smoothness of the spatial transformation s, and point correspondences between image pixels (a vector field c) are allowed to have some error.

Given a *fixed image* $F(.)$ and a *moving image* $M(.)$, the following global energy

is optimized:

$$E(c,s) = \frac{1}{\sigma_i^2}\,\mathrm{Sim}\,(F, M\circ c) + \frac{1}{\sigma_x^2}\,\mathrm{dist}\,(s,c)^2 + \frac{1}{\sigma_T^2}\,\mathrm{Reg}\,(s)\,, \qquad (1)$$

where σ_i accounts for the noise on the image intensity, σ_x accounts for spatial uncertainty on the correspondences and σ_T controls the amount of regularization. Classically, $\mathrm{Sim}\,(F, M\circ s) = \frac{1}{2}\,\|F - M\circ s\|^2$, $\mathrm{dist}\,(s,c) = \|c-s\|$ and $\mathrm{Reg}\,(s) = \|\nabla s\|^2$ but the regularization can also be modified to handle fluid-like constraints.

Within this framework, the Demons registration can be explained as an alternate optimization over s and c. The optimization is performed within the complete space of dense non-rigid transformations by taking a series of additive steps, $s \leftarrow s + \boldsymbol{u}$. In constrast, the diffeomorphic demons algorithm optimizes $E(c,s)$ over a space of diffeomorphisms. This is done in (**?**) by using an intrinsic update step, $s \leftarrow s \circ \exp(\boldsymbol{u})$, on the group of diffeomorphisms.

(dof $\simeq 3 \times$#voxels = 21million)

8.6 FNIRT: FMRIB's Nonlinear Image Registration Tool

FNIRT was developed by Andersson et al. (**?**) of the FMRIB Centre at the University of Oxford, UK. Andersson provided pre-release software for use in this study.

FNIRT is the FMRIB tool for small-displacement non-linear registration. The displacement fields are modelled as linear combinations of a basis set of splines of order two or higher, with a default of three (cubic splines). Regularisation of the field is based on bending energy (default) or membrane energy (the prior default used in this study), with optimization by multi-scale Levenberg-Marquardt minimization. The registration is initialized and run to convergence with sub-sampled images, a field of low resolution and a high regularization weight. The images and the fields from the first step are then up-sampled, the regularization modified and it is again run to convergence. This is repeated until the required warp resolution and level of regularization is achieved. After each resolution step the field is projected back onto the space of fields with Jacobians within a predefined range (**?**).

The important and unique aspects of FNIRT are with respect to its cost function. Normally when using a sum-of-squares (SOS) cost function one minimizes the difference between a warped source image and some target. FNIRT minimizes the SOS between a warped source image and the expectation of an image in target space. This expectation is based on an actual

image in target space and an intensity model. The intensity model will have some set of parameters that is determined along with the warp parameters as part of the optimization. This model can be very simple, for example just a linear scaling of the target image, in which case the parameters will be a single scale factor. The model can also be quite complex, if for example one wants to model a spatially varying flip-angle (which is a real problem with high-field scanners) it will consist of some set (e.g. 5) fields. Each of these fields is modeled as a linear combination of some basis set (also splines) and will consist of a few thousand parameters. For each voxel the expected intensity would then be a 5th-order polynomial of the intensity in the target image, where the coefficients are given by the values of the five fields at that voxel. This way it is possible to accurately model an image where (due to inhomogeneities) one area is strongly T1-weighted and another area has a strong T2-component. There is a set of different intensity models in FNIRT and the decision of which to use is made based on the properties of the two images one attempts to match.

(dof $\simeq$ 30,000)

8.7 IRTK: Image Registration Toolkit

IRTK was developed by Rueckert et al. (**???**) of Imperial College, UK. The executables are available for download at `http://www.doc.ic.ac.uk/~dr/software/`.

IRTK uses a combined transformation $\mathbf{T}$ which consists of a global transformation and a local transformation:

$$\mathbf{T}(\mathbf{x}) = \mathbf{T}_{global}(\mathbf{x}) + \mathbf{T}_{local}(\mathbf{x}) \tag{2}$$

The global transformation describes the overall differences between the two subjects and is represented by an affine transformation. The local transformation describes any local deformation required to match the anatomies of the subjects. IRTK uses a free-form deformation (FFD) model based on B-splines. The basic idea of FFDs is to deform an object by manipulating an underlying mesh of control points. The resulting deformation controls the shape of the 3-D object and can be written as the 3-D tensor product of the familiar 1-D cubic B-splines,

$$\mathbf{T}_{local}(\mathbf{x}) = \sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{n=0}^{3} B_l(u)B_m(v)B_n(w)\mathbf{c}_{i+l,j+m,k+n} \tag{3}$$

where $\mathbf{c}$ denotes a $m_x \times m_y \times m_z$ lattice of control points which parameterise the

free-form deformation, i, j, k denote the indices of the control points and u, v, w correspond to the relative positions of $\mathbf{x}$ in lattice coordinates. The lattice of control points is defined as a grid with uniform spacing which is placed on the underlying reference image. The optimal transformation is found using a gradient descent minimisation of a cost function associated with the global transformation parameters as well as the local transformation parameters. The cost function comprises two competing goals: The first term represents the cost associated with the voxel-based similarity measure, in this case normalised mutual information (**?**), while the second term corresponds to a regularization term which constrains the transformation to be smooth.

In the current study no regularization was used. This will most likely not affect results evaluated with overlap measures, but may affect studies interested in folding in the deformation fields. A penalty term for folding in the transformation is described in (**?**).

(dof $\simeq 3 \times$#control points = 1,422,843)

8.8 JRD-fluid: Jensen-Rényi Divergence fluid

JRD-fluid was developed by Chiang et al. (**?**) at LONI, UCLA, CA. Chiang provided the command options for the executables and revised the resampling code to include an option for nearest-neighbor interpolation. All registrations were run on LONI's servers.

JRD-fluid is based on an information-theoretic measure, the Jensen-Rényi divergence. JRD is derived from the joint histogram of two images. Using variational calculus methods, the driving forces are defined throughout the deforming image to maximize the JRD between it and the target image. A viscous fluid regularizer was applied to guarantee diffeomorphic (i.e., smooth, one-to-one) deformation mappings. The resulting partial differential equation was solved iteratively by convolving the applied force field with the Green's function of the linear differential operator.

(dof $\simeq 128^3$ = 2,097,152)

8.9 ROMEO: Robust multigrid elastic registration based on optical flow

ROMEO was developed by Hellier et al. (**?**) at INRIA Rennes, France. Hellier provided the executables for this study.

The ROMEO registration method expresses the registration process as the

minimization of a cost function depending on two terms: an optical flow-based similarity measure and a regularization term. The optical flow hypothesis, introduced by Horn and Schunck (**?**), assumes that the luminance of a physical point does not change when the point moves with the flow:

$$f(s + \mathbf{w}_s, t_1) - f(s, t_2) = 0 \tag{4}$$

where s is a voxel of the volume, t_1 and t_2 are the indexes of the volumes (temporal indexes for a dynamic acquisition, indexes in a database for multi-subject registration), f is the luminance function and $\mathbf{w}$ the expected 3-D displacement field.

Generally, a linear expansion of this equation is preferred: $\nabla f(s,t) \cdot \mathbf{w}_s + f_t(s,t) = 0$, where $\nabla f(s,t)$ stands for the spatial gradient of luminance and $f_t(s,t)$ is the voxelwise difference between the two volumes. The resulting set of undetermined equations has to be complemented with some prior on the deformation field. This prior is defined according to the quadratic difference of the deformation field computed between neighbours. Using an energy-based framework the regularization problem may be formulated as the minimization of the following cost function:

$$U(\mathbf{w}; f) = \sum_{s \in S} [\nabla f(s,t) \cdot \mathbf{w}_s + f_t(s,t)]^2 + \alpha \sum_{<s,r> \in \mathcal{C}} ||\mathbf{w}_s - \mathbf{w}_r||^2 \tag{5}$$

where S is the voxel lattice, $\mathcal{C}$ is the set of neighboring pairs w.r.t. a given neighborhood system $\mathcal{V}$ on S ($< s, r > \in \mathcal{C} \Leftrightarrow s \in \mathcal{V}(r)$), and α controls the balance between the two energy terms. The first term is the linear expansion of the luminance conservation equation and represents the interaction between the field and the data. The second term is the smoothness constraint. In order to cope with large displacements, an incremental multi-resolution procedure is used to construct a pyramid of volumes by successive Gaussian blurring and subsampling.

(dof $\simeq$ 2 million for 300,000 rigid and 50,000 affine estimates)

8.10 SICLE: Small-deformation, Inverse-Consistent, Linear-Elastic image registration

SICLE was developed by Christensen et al. (**???**) at the University of Iowa. Song provided the executables and helped to install the software.

SICLE is based on the principle of jointly estimating the forward h and reverse g transformations between two images while minimizing the inverse consistency error $||h(x) - g^{-1}(x)|| + ||g(x) - h^{-1}(x)||$. Ideally, the forward transformation h from image T to S and the reverse transformation g from

S to T should be uniquely determined and should be inverses of one another. However, estimating h and g independently as with most uni-directional image registration algorithms rarely results in an inverse consistent set of transformations due to a large number of local minima. Jointly estimating the forward and reverse transformations provides additional correspondence information helping to minimize correspondence errors.

The SICLE image registration algorithm iteratively minimizes the following cost function

$$\begin{aligned} C =& \sigma \int_{\Omega} |T_i(h_{i,j}(x)) - T_j(x)|^2 + |T_j(h_{j,i}(x)) - T_i(x)|^2 dx \\ &] + \rho \int_{\Omega} ||\mathcal{L}u_{i,j}(x)||^2 + ||\mathcal{L}u_{j,i}(x)||^2 dx \\ & + \chi \int_{\Omega} ||u_{i,j}(x) - \tilde{u}_{j,i}(x)||^2 + ||u_{j,i}(x) - \tilde{u}_{i,j}(x)|| dx. \end{aligned} \tag{6}$$

where the parameters σ, ρ, and χ are weighting constants used to vary the influence of each term of the cost function and Ω is the image domain. The intensity of the images are normalized between 0 and 1. The first integral of the cost function defines the correspondence (squared intensity difference) between the deformed template and target images and between the deformed target and template images, respectively. The second integral is used to regularize the forward and reverse displacement fields $u_{i,j}$ and $u_{j,i}$ respectively, and is minimized when the forward and reverse displacement fields satisfy the properties of the linear elastic model. In the linear elastic model, the linear differential operator $\mathcal{L}u(x) = \alpha\nabla^2 u(x) + \beta\nabla(\nabla \cdot u(x)) + \gamma u(x)$ penalizes large second and cross derivatives in the x, y, and z directions in the displacement fields. The constant weighting parameters α, β, and γ are used to vary the influence of the terms of $\mathcal{L}u(x)$. The third integral is called the inverse consistency constraint and is minimized when the forward and reverse transformations $h_{i,j}$ and $h_{j,i}$, respectively, are inverses of each other.

The cost function in Equation 6 is minimized using the gradient descent method described in (**?**) where each component of the displacement field is parameterized in terms of a 3-D Fourier series. A multi-resolution approach in the frequency and spatial domains is taken to estimate the Fourier Series coefficients by first estimating the low frequency components then increasing the number of harmonics as the estimation progresses.

(dof = 7,986 for 10 harmonics)

8.11 SPM5: Statistical Parametric Mapping

SPM5 registration algorithms were developed by Ashburner et al.(**???**) at the Functional Imaging Laboratory, UK. Normalize, Unified Segmentation, and the DARTEL Toolbox are part of the SPM5 package and may be downloaded from `http://www.fil.ion.ucl.ac.uk/spm/software/spm5/`. Ashburner provided Matlab scripts for composing the Normalization, Unified Segmentation and DARTEL Toolbox transforms.

Five methods were evaluated from the SPM5 package: regular and "SPM2-type" Normalization, Unified Segmentation, and regular and pairwise implementations of the DARTEL Toolbox. It is expected that the SPM user will register original brain images (non-skull-stripped and in their native space) to idealized templates. The template images supplied with SPM5 conform to the space defined by the ICBM, NIH P-20 project, and approximate that of the space described in the atlas of Talairach and Tournoux (Talairach and Tournoux, 1988). In this study, the "SPM2-type" Normalization and the pairwise DARTEL were the only SPM methods that were applied in the same manner as all of the other non-SPM algorithms, that is, by directly registering one skull-stripped brain image in MNI space to another (rather than to one another via a coregistration template).

SPM5 regular and "SPM2-type" Normalization (**?**)

The algorithms work by minimizing the sum of squares difference between the image which is to be normalized and a linear combination of one or more template images (in the case of "SPM2-type" Normalization, a single target image). The first step of the normalization is to determine the optimum 12-parameter affine transformation. A Bayesian framework is used, such that the registration searches for the solution that maximizes the *a posteriori* probability of it being correct. That is, it maximizes the product of the likelihood function (derived from the residual squared difference) and the prior function (which is based on the probability of obtaining a particular set of zooms and shears).

The affine registration is followed by estimating nonlinear deformations, whereby the deformations are defined by a linear combination of 3-D discrete cosine transform (DCT) basis functions. The parameters represent coefficients of the deformations in three orthogonal directions. The matching involves simultaneously minimizing the bending energies of the deformation fields and the residual squared difference.

(dof $\simeq$ 1,000)

SPM5 Unified Segmentation (**?**)

This approach uses a probabilistic generative model that combines image registration, tissue classification, and bias correction. The log-likelihood objective function is based on a mixture of Gaussians, and is extended to incorporate a smooth intensity variation and nonlinear registration with tissue probability maps. A small-deformation registration model is used, which is parameterised by a linear combination of around 1,000 cosine transform basis functions.

In addition to simple inter-subject registration, the Unified Segmentation approach also performs classification of brain tissues into gray and white matter, as well as bias correction and rudimentary skull-stripping.

(dof $\simeq$ 1,000)

SPM5 DARTEL Toolbox: Diffeomorphic Anatomical Registration using Exponentiated Lie algebra (**?**)

This approach is an extension of the approach described by Ashburner (**?**), and is intended to register tissue class images from multiple subjects with a common template. The tissue class images are typically gray and white matter, which have been extracted using the SPM5 Unified Segmentation approach. The generative model assumes that the template encodes the mean of a multinomial distribution. Pre-computed templates can be used, but the toolbox also allows them to be iteratively generated from a population of subjects using a "congealing" approach. Nonlinear registration is considered as a local optimization problem, which is done using a Levenberg-Marquardt strategy. The necessary matrix solutions are obtained in reasonable time using a multi-grid method. A constant Eulerian velocity framework is used, which allows a rapid scaling and squaring method to be used in the computations.

(dof = 121×145×121×3 = 6,368,835)

8.12 SyN: Symmetric Normalization

SyN was developed by Brian Avants et al. (**?**) at the University of Pennsylvania. Brian Avants provided the executables for a beta version for this study. Several implementation improvements have been made since the algorithm was evaluated in this work. SyN is available in the ANTS toolkit; a new release with tutorial is available at `http://www.picsl.upenn.edu/ANTS/`.

The symmetric normalization (SyN) methodology uses a symmetric parameterization of the shortest path of diffeomorphisms connecting two neuroanatomical configurations. The SyN formulation uses a bidirectional

gradient descent optimization which gives results that are unbiased with respect to the input images. SyN also provides forward and inverse continuum mappings that are consistent within the discrete domain and enables both large and subtle deformations to be captured. Specific performance characteristics depend upon the range of similarity metrics chosen for the study and the velocity field regularization. The current study uses Gaussian smoothing of the velocity field and a gradient-based optimization of an approximate cross-correlation (CC) similarity metric with CC evaluated in a window of size 5x5x5 voxels. A variety of other similarity measures are available, including robust optical flow, mutual information and additional correlation measures.

(dof $\simeq 4 \times$#voxels = 28million)

References

Allen, J. S., Bruss, J., Mehta, S., Grabowski, T., Brown, C. K., Damasio, H., 2008. Effects of spatial transformation on regional brain volume estimates. NeuroImage 42, 535–547.

Bechhofer, R. E., Mar. 1954. A single-sample multiple decision procedure for ranking means of normal populations with known variances. The Annals of Mathematical Statistics 25, 16–39.

Beg, M. F., Miller, M. I., Trouvé, A., Younes, L., Feb. 2005. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. International Journal of Computer Vision 61, 139–157.

Cachier, P., Bardinet, E., Dormont, D., Pennec, X., Ayache, N., 2003. Iconic feature based nonrigid registration: the PASHA algorithm. Computer Vision and Image Understanding 89, 272–298.

Caviness, V., Meyer, J., Makris, N., Kennedy, D., 1996. MRI-based topographic parcellation of human neocortex: an anatomically specified method with estimate of reliability. Journal of Cognitive Neuroscience 8, 566–587.

Clouchoux, C., Coulon, O., Rivière, D., Cachia, A., Mangin, J.-F., Régis, J., 2005. Anatomically constrained surface parameterization for cortical localization. Medical Image Computing and Computer-Assisted Intervention: MICCAI 2005 3750, 344–351.

Cointepas, Y., Mangin, J.-F., Garnero, L., Poline, J.-B., Benali, H., Jun. 2001. BrainVISA: Software platform for visualization and analysis of multi-modality brain data. NeuroImage 13, 98.

Crum, W. R., Camara, O., Rueckert, D., Bhatia, K. K., Jenkinson, M., Hill, D. L. G., 2005. Generalised overlap measures for assessment of pairwise and groupwise image registration and segmentation. Medical Image Computing and Computer-Assisted Intervention: MICCAI 2005 3749, 99–106.

Crum, W. R., Griffin, L. D., Hill, D. L. G., Hawkes, D. J., 2003. Zen and the art of medical image registration: correspondence, homology, and quality. NeuroImage 20, 1425–1437.

Crum, W. R., Rueckert, D., Jenkinson, M., Kennedy, D., Smith, S. M., 2004. A framework for detailed objective comparison of non-rigid registration algorithms in neuroimaging. Medical Image Computing and Computer-Assisted Intervention: MICCAI 2004 3216, 679–686.

Dale, A. M., Fischl, B., Sereno, M. I., Feb. 1999. Cortical surface-based analysis I: Segmentation and surface reconstruction. NeuroImage 9, 179–194.

Drury, H. A., Essen, D. C. V., Anderson, C. H., Lee, C. W., Coogan, T. A., Lewis, J. W., 1996. Computerized mappings of the cerebral cortex: A multiresolution flattening method and a surface-based coordinate system. Journal of Cognitive Neuroscience 8, 1–28.

Essen, D. C. V., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., Anderson, C. H., 2001. An integrated software suite for surface-based analyses of cerebral cortex. Journal of the American Medical Informatics Association :

JAMIA 8, 443–59.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., Peters, T. M., Oct. 1993. 3D statistical neuroanatomical models from 305 MRI volumes. Nuclear Science Symposium and Medical Imaging Conference 3, 1813–1817.

Evans, A. C., Collins, D. L., Milner, B., 1992. An MRI-based stereotactic brain atlas from 300 young normal subjects. Proc. of the 22nd Symposium of the Society for Neuroscience, Anaheim 408.

Fiez, J. A., Damasio, H., Grabowski, T. J., 2000. Lesion segmentation and manual warping to a reference brain: intra- and interobserver reliability. Human Brain Mapping 9, 192–211.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A. M., 2002. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341–355.

Fischl, B., Sereno, M. I., Tootell, R. B. H., Dale, A. M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Human Brain Mapping 8, 272–284.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A. M., 2004. Automatically parcellating the human cerebral cortex. Cerebral Cortex 14, 11–22.

Friston, K. J., Ashburner, J., Poline, J. B., Frith, C. D., Heather, J. D., Frackowiak, R. S. J., 1995. Spatial registration and normalization of images. Human Brain Mapping 2, 165–189.

Gee, J. C., Reivich, M., Bajcsy, R., 1993. Elastically deforming 3D atlas to match anatomical brain images. Journal of Computer Assisted Tomography 17, 225–236.

Gerig, G., Jomier, M., Chakos, M., 2001. Valmet: A new validation tool for assessing and improving 3D object segmentation. Lecture Notes in Computer Science 2208, 516–524.

Gholipour, A., Kehtarnavaz, N., Briggs, R., Devous, M., Gopinath, K., 2007. Brain functional localization: A survey of image registration techniques. IEEE Transactions on Medical Imaging 26, 427–451.

Grachev, I. D., Berdichevsky, D., Rauch, S. L., Heckers, S., Alpert, N. M., 1998. Anatomic landmark-based method for assessment of intersubject image registration techniques: Woods vs Talairach [abstract]. 8th Annual Meeting of the Organization for Human Brain Mapping, Brighton, England.

Grachev, I. D., Berdichevsky, D., Rauch, S. L., Heckers, S., Kennedy, D. N., Caviness, V. S., Alpert, N. M., 1999. A method for assessing the accuracy of intersubject registration of the human brain using anatomic landmarks. NeuroImage 9, 250–268.

Heckemann, R. A., Hajnal, J. V., Aljabar, P., Rueckert, D., Hammers, A., Oct. 2006. Automatic anatomical brain MRI segmentation combining label

propagation and decision fusion. NeuroImage 33, 115–126.
Hellier, P., 2003. Consistent intensity correction of MR images. International Conference on Image Processing: ICIP 2003 1, 1109–1112.
Hellier, P., Ashburner, J., Corouge, I., Barillot, C., Friston, K., 2002. Inter subject registration of functional and anatomical data using SPM. Medical Image Computing and Computer-Assisted Intervention: MICCAI 2002 2489, 590–597.
Hellier, P., Barillot, C., Corouge, I., Gibaud, B., Goualher, G. L., Collins, D. L., Evans, A., Malandain, G., Ayache, N., Christensen, G. E., Johnson, H. J., Sep. 2003. Retrospective evaluation of intersubject brain registration. IEEE transactions on medical imaging 22, 1120–30.
Hellier, P., Barillot, C., Corouge, I., Gibaud, B., Goualher, G. L., Collins, L., Evans, A., Malandain, G., Ayache, N., 2001. Retrospective evaluation of inter-subject brain registration. Lecture Notes in Computer Science 2208, 258–266.
Jaccard, P., 1912. The distribution of flora in the alpine zone. The New Phytologist 11, 37–50.
Jannin, Grova, Maurer, 2006. Model for defining and reporting reference-based validation protocols in medical image processing. International Journal of Computer Assisted Radiology and Surgery 1, 63–73.
Kittler, J., Hatef, M., Duin, R. P. W., Matas, J., 1998. On combining classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence 20, 226–239.
Klein, A., Hirsch, J., 2005. Mindboggle: a scatterbrained approach to automate brain labeling. NeuroImage 24(2), 261–280.
Klein, A., Mensh, B., Ghosh, S., Tourville, J., Hirsch, J., 2005. Mindboggle: Automated brain labeling with multiple atlases. BMC Medical Imaging 5:7.
Mandl, R., Jansma, J., Slagter, H., Collins, D., Ramsey, N., 2000. On the validity of associating stereotactic coordinates with anatomical nomenclature. Human Brain Mapping: HBM 2000: S539.
Menke, J., Martinez, T., 2004. Using permutations instead of student's t distribution for p-values in paired-difference algorithm comparisons. In: Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on. Vol. 2. pp. 1331–1335 vol.2.
Nieto-Castanon, A., Ghosh, S. S., Tourville, J. A., Guenther, F. H., 2003. Region of interest based analysis of functional imaging data. NeuroImage 19, 1303–1316.
Paglieroni, D. W., 1992. Distance transforms: properties and machine vision applications. CVGIP: Graph. Models Image Process. 54, 56–74.
Rogelj, P., Kovacic, S., Gee, J., 2002. Validation of a nonrigid registration algorithm for multimodal data. Proc. SPIE 4684, 299–307.
Rohlfing, T., Russakoff, D. B., Maurer, C. R., 2004. Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. IEEE Transactions on Medical Imaging 23, 983–994.

Roland, P., Geyer, S., Amunts, K., Schormann, T., Schleicher, A., Malikovic, A., Zilles, K., 1997. Cytoarchitectural maps of the human brain in standard anatomical space. Human Brain Mapping 5, 222–227.

Rosenfeld, A., Pfaltz, J. L., 1966. Sequential operations in digital picture processing. J. ACM 13, 471–494.

Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., Poldrack, R. A., Bilder, R. M., Toga, A. W., Feb. 2008. Construction of a 3D probabilistic atlas of human cortical structures. NeuroImage 39, 1064–1080.

Shen, D., Davatzikos, C., Nov. 2002. Hammer: hierarchical attribute matching mechanism for elastic registration. IEEE transactions on medical imaging 21, 1421–39.

Talairach, J., Szikla, G., 1967. Atlas d'anatomie stereotaxique du telencephale: etudes anatomo-radiologiques. Masson & Cie, Paris.

Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain. Thieme Medical Publishers, New York.

Tourville, J., Guenther, F., 2003. A cortical and cerebellar parcellation system for speech studies. Boston University Technical Reports CAS/CNS-03-022.

Towle, V. L., Khorasani, L., Uftring, S., Pelizzari, C., Erickson, R. K., Spire, J.-P., Hoffmann, K., Chu, D., Scherg, M., Jul. 2003. Noninvasive identification of human central sulcus: a comparison of gyral morphology, functional mri, dipole localization, and direct cortical mapping. NeuroImage 19, 684–97.

Warfield, S., Zou, K., Wells, W., 2004. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. IEEE Transactions on Medical Imaging 23, 903–921.

Xiong, J., Rao, S., Jerabek, P., Zamarripa, F., Woldorff, M., Lancaster, J., Fox, P. T., Sep. 2000. Intersubject variability in cortical activations during a complex language task. NeuroImage 12, 326–339.

Yassa, M. A., Stark, C. E. L., 2009. A quantitative evaluation of cross-participant registration techniques for MRI studies of the medial temporal lobe. NeuroImage 44, 319–327.

Zijdenbos, A., Dawant, B., Margolin, R., Palmer, A., 1994. Morphometric analysis of white matter lesions in mr images: method and validation. IEEE Transactions on Medical Imaging 13, 716–724.