A New Method for Assessing PET-MRI Coregistration

Christine DeLorenzo*[a], Arno Klein[a], Arthur Mikhno[a] , Neil Gray[a], Francesca Zanderigo[b], J. John Mann[a,b], Ramin V. Parsey[a,b]

[a]Department of Psychiatry, Columbia University, 1051 Riverside Drive, New York, New York, 10032; [b]Department of Molecular Imaging and Neuropathology, New York State Psychiatric Institute, 1051 Riverside Drive, New York, New York 10032

ABSTRACT

Positron emission tomography (PET) images are acquired for many purposes, from diagnostic assessment to aiding in the development of novel therapies. Whatever the intended use, it is often necessary to distinguish between different anatomical regions within these images. Because of this, magnetic resonance images (MRIs) are generally acquired to provide an anatomical reference. This reference will only be accurate if the PET image is properly coregistered to the MRI; yet currently, a method to evaluate PET-MRI coregistration accuracy does not exist. This problem is compounded by the fact that two visually indistinguishable coregistration results can produce estimates of ligand binding that vary significantly. Therefore, the focus of this work was to develop a method that can evaluate coregistration performance based on measured ligand binding within certain regions of the coregistered PET image. The evaluation method is based on the premise that a more accurate coregistration will result in higher ligand binding in certain anatomical regions defined by the MRI. This fully automated method was able to assess coregistration results within the variance of an expert manual rater and shows promise as a possible coregistration cost function.

Keywords: Registration, Functional imaging, Atlases

1. INTRODUCTION

Positron emission tomography (PET) is a noninvasive imaging modality that can reveal functional activity in the brain. In many PET studies, it is necessary to quantify this activity in a specific anatomical region of interest (ROI) or compare activity between ROIs. However, the limited resolution and lack of spatial information inherent in PET images can make localization of these anatomical ROIs difficult or impossible. Because of this, conventional magnetic resonance images (MRIs) are usually acquired in conjunction with PET. PET images can then be spatially normalized, or coregistered, to the MRI, facilitating ROI analysis.

Though this normalization can be performed using standard software such as SPM (Wellcome Trust Centre for Neuroimaging, London, UK), the resulting coregistration accuracy can be difficult to assess. Many normalization routines attempt to optimize variants of mutual information (MI) to improve image transformation estimates. MI measures the degree to which knowledge of the source image affects the uncertainty in knowing the target [1]. It is normally a robust measure of image overlap; however it may not be an accurate predictor of coregistration accuracy when the source and target images have limited dependence, which can happen depending on the PET ligand binding pattern. Therefore, a registration that yields a minimum cost based on MI cost function may still not produce an optimal transformation. The reduced amount of spatial information present in PET images also makes visual assessment of the coregistration problematic.

Assessment of PET-MRI coregistration is therefore a difficult problem. And, unfortunately, the resulting quantitative information obtained from PET images can be highly sensitive to this coregistration accuracy. An illustration of this using the PET radioligand *N*-(3-iodoprop-2*E*-enyl)-2β-carbomethoxy-3β-(4-methylphenyl)nortropane (PE2I) can be seen in Figure 1. The specific binding of PE2I creates images that contain limited spatial information outside the known regions of high binding – caudate, putamen and ventral striatum. Because these high binding regions are the only easily distinguishable anatomical landmarks, similar coregistration results can appear virtually indistinguishable. However, estimates of ligand volume of distribution (V_T) from these similar coregistrations differed by greater than 20%.

*cd2415@columbia.edu; phone 1 212 543-6103; fax 1 212 543-6017

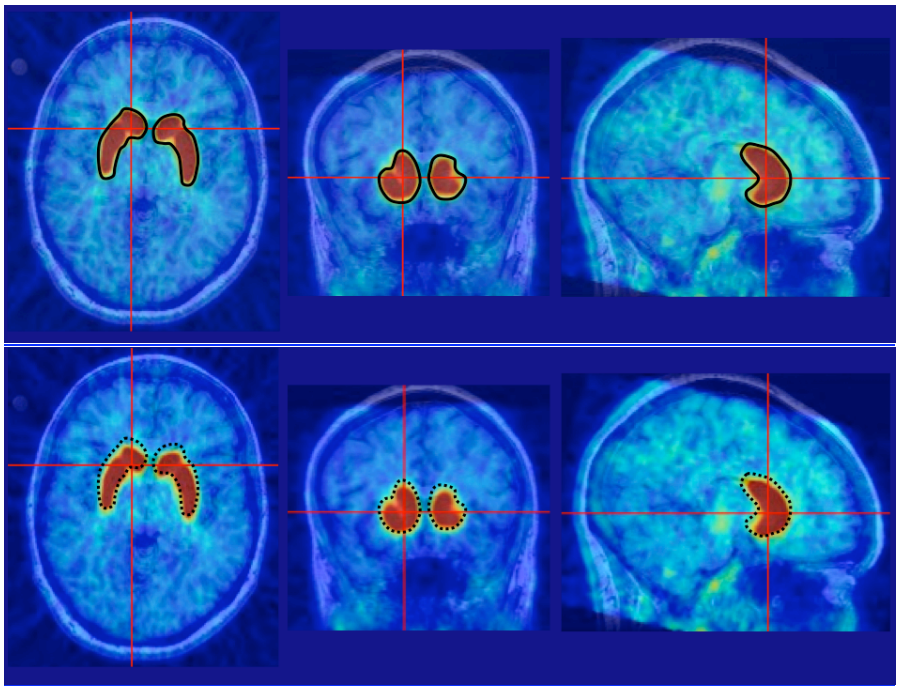

Fig. 1. Two possible coregistration solutions for PE2I are shown. Coregistered PET images are overlaid on the MRI. The same axial, sagittal and coronal (left, middle and right, respectively) image slices are shown in top and bottom panes. High binding regions are outlined in black on the top image. For reference, these outlines are also shown on the bottom (dashed line). Though the images appear similar, the top set resulted in a 20% higher ventral striatum V_T.

2. METHODS

In this work, we developed an automated system for evaluating PET-MRI coregistration. The hypothesis behind the evaluation procedure is as follows: If PET images are inaccurately transformed into MRI space, when MRI-defined anatomical structures (grey matter/cerebrum/ROIs) are overlaid on the PET images, the ligand binding within those structures will be lower than if the registration were accurate. For example, a misregistration can cause the cerebrum, as defined in the MRI, to overlap nonbrain regions of the PET image. Similarly, the regions defined in the MRI as grey matter (GM) or high-binding ROIs could overlap white matter or adjacent lower-binding ROIs. This would result in an underestimation of ligand binding within anatomical structures. This binding measure could therefore be used to select the most accurate coregistration from a set of possible coregistrations, as we do in this study, or possibly be used as a cost function to guide the coregistration procedure.

2.1 Image Acquisition

Seven healthy volunteers were used for this study. All subjects underwent two identical scans, test–retest, on the same day, separated by an approximately one-hour break. During each scan, arterial and venous catheters were placed for blood sampling and radioisotope injection, respectively. To prevent movement, individual polyurethane molds (Soule Medical, Tampa, FL, USA) were poured and fixed around the subject's head. Positron Emission Tomography was performed with an ECAT HR+ scanner (Siemens/CTI, Knoxville, TN, USA). A 10-min transmission scan was obtained before radioligand injection. At the end of the transmission scan, between 15.13 and 18.48 mCi (mean: 17.02 mCi; standard deviation: 1.10 mCi) of [^{11}C]PE2I were administered intravenously as a bolus over 30 seconds. Emission data were collected in three-dimensional mode for 120 minutes using 21 frames of increasing duration: 3 at 20 seconds, 3 at 1 minute, 3 at 2 minutes, 2 at 5 minutes, and 10 at 10 minutes. Images were reconstructed, using attenuation correction from the transmission data, to a 128 by 128 matrix (pixel size: 1.72 by 1.72 mm). A model-based method was used to correct scatter [2]. A Shepp 0.5 (2.5 mm in full width at half maximum) filter was used for the reconstruction and estimated image. The Z filter was all-pass 0.4 (2.0 mm in full width at half maximum), and the zoom factor was 4.0, leading to a final image resolution of 5.1 mm in full width at half maximum at the center of the field of view [3].

Magnetic resonance images were acquired on a 3-T Signa Advantage system (GE Healthcare, Waukesha, WI, USA), as described in [4]. The final voxel size was 1.02 x 1.02 x 1.00 mm, with an acquisition time of 11 minutes.

2.2 Input Function Measurement

An automated blood sampling system was used to collect arterial samples every ten seconds for the first two minutes and every twenty seconds between minutes two and four. Following that, ten samples were collected manually at 5, 8, 10, 16, 20, 30, 40, 60, 80, 90 and 110 minutes, for a total of 30 samples. Each blood sample was then centrifuged, and plasma supernatant collected in 200 mL aliquots, from which the radioactivity was measured in a well counter. A high pressure liquid chromatography assay of seven of the collected blood samples (at 2, 5, 10, 30, 50, 80 and 100 mins) was used to established unmetabolized parent compound levels. These unmetabolized parent fraction levels were fit with a Hills function [5]. To account for metabolism and radioactive decay, the input function was calculated as the product of the interpolated parent fraction and the total plasma counts. The calculated input function was then fit as the combination of a second order polynomial and the sum of three exponentials, describing the function before and after the peak, respectively.

2.3 Coregistration Procedure

To correct for patient motion during the PET scan, we registered the last thirteen frames of an individual PET study to the eighth frame using the FMRIB linear image registration tool (FLIRT, version 5.4.2, FMRIB Image Analysis Group, Oxford, UK). We then used FLIRT to compute eight different PET- MRI rigid-body transformations with a mutual information cost function, six degrees of freedom and trilinear interpolation. The eight coregistrations each had a different combination of source PET volume, target MRI volume and cropping mask (Table 1). These represent the range of possibilities often used for PET-MRI coregistration in ligand studies.

To obtain the various target images and masks, the MRI volumes needed to be skull-stripped and segmented. We used FMRIB's Brain Extraction Tool (BET) to perform skull-stripping and the SPM5 (Wellcome Trust Centre for Neuroimaging, London, UK) spm_segment function to obtain a grey and white matter segmentation.

Table 1. Coregistration options. Eight different coregistrations were tested. In the table, "perfusion frames" refers to the mean of the first seven frames in which the blood perfusion is visible.

coregistration number	PET frames	target MRI	mask
1	mean of 8 - 20	skull-stripped	none
2	mean of 8 - 20	complete MRI	grey & white matter
3	mean of 8 - 20	skull-stripped	grey & white matter
4	mean of 8 - 20	grey & white matter	grey & white matter
5	perfusion frames	grey & white matter	grey & white matter
6	perfusion frames	complete MRI	grey & white matter
7	mean of 8 - 20	complete MRI	none
8	mean of 8 - 20	grey matter	none

2.4 Parameter Estimation

To evaluate coregistration accuracy, we created parameter images in which volume of distribution (V_T) or binding potential (BP_{ND}) of the ligand was modeled at each voxel. In this case, V_T reflects the volume of blood (or plasma) it would take to account for the amount of drug in the part of the brain represented by a particular voxel [6]. Using

graphical analysis, V_T can be readily calculated as the slope of the linear part of a $\frac{\int_0^t ROI(t')dt'}{ROI(t)}$ versus $\frac{\int_0^t C_p(t')dt'}{ROI(t)}$ plot, where ROI(t) represents the total activity in the voxel of interest at a certain time, t, and C_p is the concentration of the ligand in the plasma. This method, proposed by Logan *et al.* [7], is computationally efficient and fast; however, measurements of V_T require blood sampling (as described in section 2.2). In contrast, BP_{ND}, the ratio of specifically to nonspecifically bound ligand, can be calculated using a reference region, which is assumed to be devoid of specific binding. BP_{ND} can be calculated by replacing $C_p(t)$ in the graphical analysis with $ROI_{ref}(t)$, the concentration of the ligand in the reference tissue.

Once V_T and BP_{ND} parameter images were created (in PET space), we applied the eight different PET-MRI transformations to transform the parameter images into MRI space. We extracted summed V_T or BP_{ND} values in the high binding ROIs (caudate, putamen, ventral striatum and midbrain), grey matter or over the entire cerebrum. The ROIs were obtained in an automated fashion using an in-house probabilistic brain atlas that was normalized to the MRI using the SPM5 spm_normalise function. Cerebrum was determined by thresholding the skull-stripped MRI image. The best coregistration was defined as that which resulted in the highest summed V_T or BP_{ND} value for each of the three regions. This process is outlined in Figure 2.

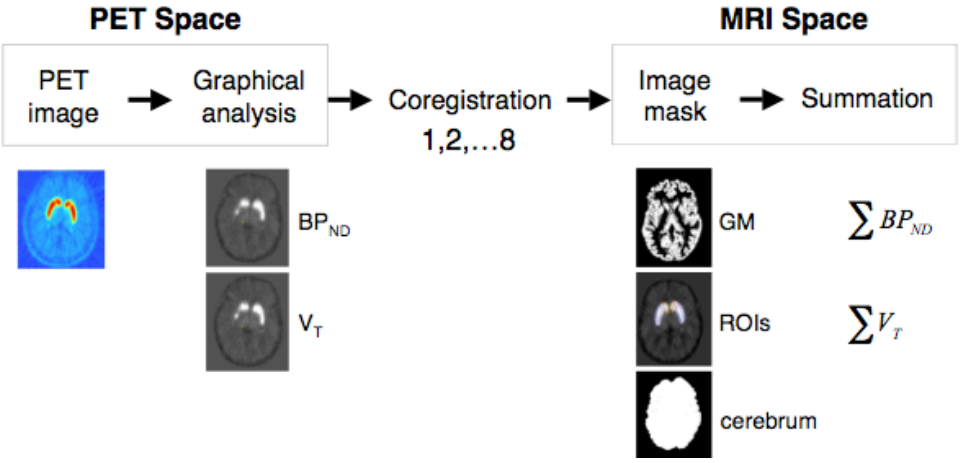

Fig. 2. An outline of the method used to evaluate coregistration accuracy. BP_{ND} and V_T images were created from the PET image, coregistered using one of eight methods, and then the value of the parameter (BP_{ND} or V_T) was summed within an image mask.

2.5 Validation

In order to visually assess coregistration validity, we used in-house software to overlay the coregistered PET volumes on their corresponding MRI volumes, resulting in images similar to those shown in Figure 1. An expert rater then used the in-house software's interface to navigate through the slices of the superimposed images and record the coregistration that appeared most accurate. The rater evaluated all eight coregistration possibilities for the seven different PE2I PET scans. To assess the intrarater reliability, the rater was asked to select the best performing coregistration on two different occasions, separated by over one month.

Though the scope of this project involves a novel method to determine which of a set of coregistrations is most accurate, using summed BP_{ND} or V_T values in this way naturally lends itself to viewing these values as a cost function within the coregistration procedure. In order to use summed parameter values in this manner, it is important that the cost function (negative of the summed V_T/BP_{ND} values) reaches a minimum when the coregistration is correct. We tested the cost function in two ways. First, we offset the coregistered PET images considered most accurate by the expert user with linearly increasing translations and rotations, and calculated the resulting cost function value. In this test, the each

rotation or translation was applied independently. Second, we ran a similar test with jointly applied, randomly selected rotations and translations.

2.6 Application to Test-Retest Study

We calculated average BP_{ND} values for the caudate, putamen and ventral striatum for each of the seven subjects (both test and retest) and for each of the eight coregistrations (Table 1, section 2.3). We chose the best coregistration method, either based on the parameter images (as shown in Figure 2) or as the one that produced the lowest MI cost across the eight coregistration results. We also calculated the percent difference (PD, the absolute difference between test and retest values on the same subject, divided by their average) between test and retest BP_{ND} values using both coregistration choosing techniques results.

3. RESULTS

In order to compare coregistrations, we calculated the mean V_T and BP_{ND} values in areas of high binding (caudate, putamen, ventral striatum and midbrain) for the coregistrations performed on the first of the subject's two (test/restest) images (Figure 3). Two coregistrations are considered similar if they produce summed BP_{ND} (or V_T) values that are within 5% of each other for every region. Figure 3 demonstrates (with a single subject) why these coregistrations are difficult to assess visually. In this case, though the eight coregistrations produced similar summed BP_{ND} values for the caudate (except coregistration #8) and putamen (except #6), two of the coregistrations (#2 and #6) produced low BP_{ND} values for the midbrain. And, the ventral striatum BP_{ND} values were only similar for coregistrations #4 and #7. This means that even though some regions may appear to be correctly aligned, the coregistration was still inaccurate.

Aside from allowing a simple view of coregistration similarity, Figure 3 elucidates some of the problems that may occur in image matching. For example, coregistration #8 produced high values of BP_{ND} in the ventral striatum. This is most likely due to a misregistration of lower binding ventral striatum in the PET volume being matched to the higher binding caudate or putamen in the MRI (and therefore mislabeled as caudate or putamen). The caudate and putamen, in turn, overlap low binding regions, explaining the pattern for coregistration #8 in which the ventral striatum BP_{ND} values are much higher, at the expense of BP_{ND} values in the other regions.

The results of the manual evaluation are listed in the first two rows of Table 2. As in Figure 3, coregistrations that produced mean V_T or BP_{ND} values within 5% of those calculated from the rater's first choice coregistration (row 1 of Table 2) were considered as similar coregistrations and highlighted in grey. Since results obtained with the V_T parameter images were the same as those of the BP_{ND} parameter images in almost all cases, all results presented here are based on BP_{ND} values. As can be seen in Table 2, the expert rater was consistent (chose a similar coregistration) for five out of seven (71.4%) of the repeat trials.

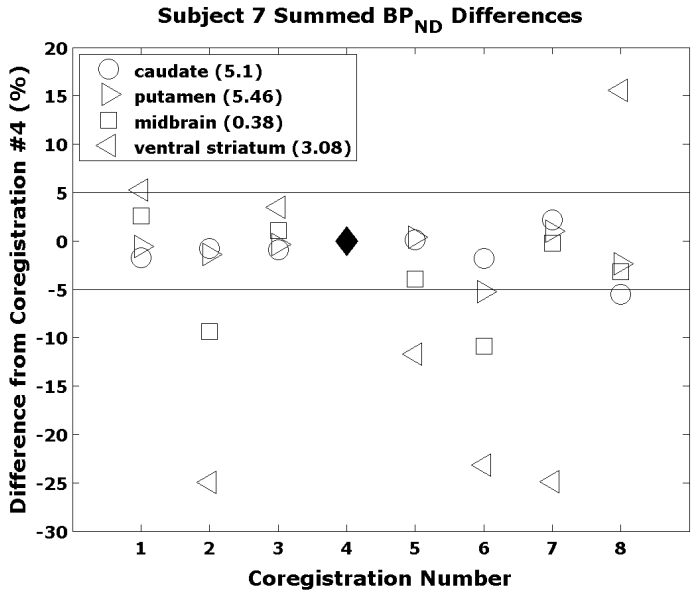

Fig. 3. Representation of the results for subject 7. The rater's first choice for best coregistration was #4 (black diamond), which resulted in mean BP_{ND} values as indicated in the legend. The mean BP_{ND} values for the other coregistration results were calculated and normalized to the legend values. Points that fall within the horizontal lines indicate regions with BP_{ND} values that differ less than 5% from coregistration #4. This plot indicates the sensitivity of small regions, such as the ventral striatum, to coregistration.

Table 2. Coregistration evaluation: manual vs. automated consistency. The first two rows indicate the best coregistration procedure as assessed by the expert user. Rows 3- 5 indicate the best coregistration as determined by the summed BP_{ND} value in the grey matter, cerebrum, or regions of interest (ROIs). The last row is the same as row 5, except the mean PET image (not the parameter image) was used. Grey shading indicates mean BP_{ND} values that are within 5% of the rater's first choice. Multiple numbers in one cell indicate coregistrations that appeared indistinguishable.

Subject:	1	2	3	4	5	6	7
Rater's first choice	2	1	1	4	6	2	4
Rater's second choice	5	6	1	2,3,4	2	2	3
BP_{ND} (GM)	1	3	1	1	4	4	3
BP_{ND} (cerebrum)	1	1	5	2	2	1	1
BP_{ND} (ROIs)	3	1	1	4	2	1	4
PET image (ROIs)	3	1	6	4	2	1	1

3.1 Manual/Automated Evaluation Comparison

Table 2 shows the results of using the summed BP_{ND} values within GM (row 3), cerebrum (row 4) or ROIs (row 5). The summed BP_{ND} values in the ROIs produced the most consistent results with the rater's first choice rankings. This method selected similar results in six out of seven (85.7%) of the trials, which was higher than the intrarater consistency. For comparison, the summed activity from the PET image (mean of all frames) was used to rank the coregistrations. It did not perform as well, indicating the importance of using BP_{ND} (or V_T) parameter images.

3.2 Validation Check

As seen in Figure 4, the summed ROI BP_{ND} cost function behaves as expected and is minimized as coregistration results become more accurate. In the case of translation in the z-direction and rotation in the x-axis, the normalized cost values reach an asymptote, meaning that, past a certain value, there is little change in the summed BP_{ND} values as translation/rotation offsets increase. This would be a problem for an optimization routine using summed BP_{ND} as a cost function because the gradient of the costs would be zero in the flat regions and it would be unclear how to minimize the function. However, it is important to note that these values are far from the correct position and orientation (> 12 mm or > 10 degrees off) and the coregistration technique will rarely be initialized with values that far from true values. To avoid this problem, the images could be roughly aligned with a standard image registration technique prior to BP_{ND}-based registration.

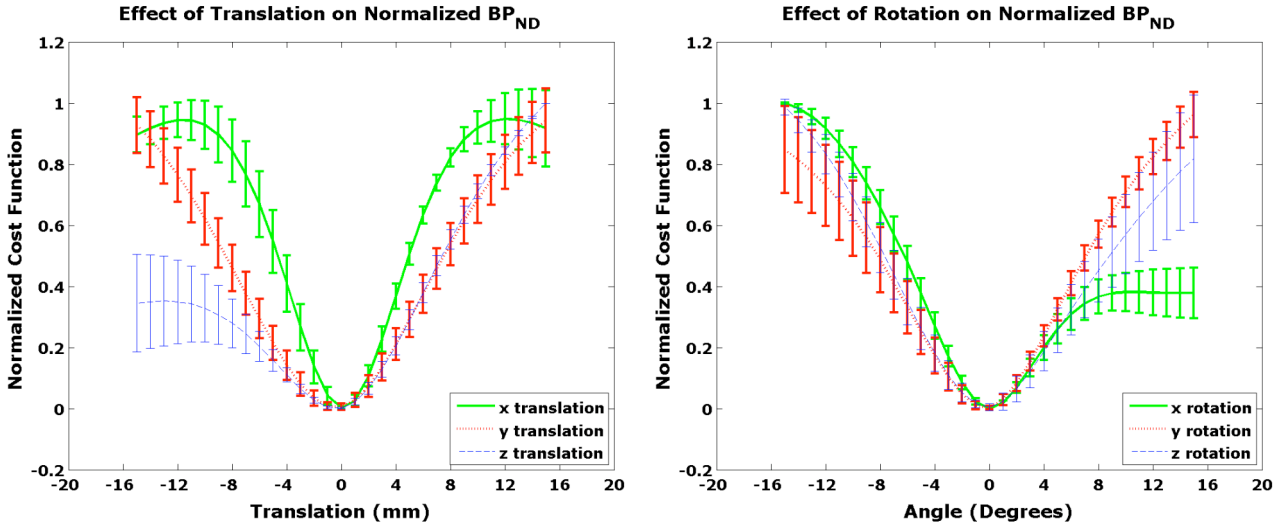

Fig. 4. Effects of translation (**Left**) or rotation (**Right**) along each axis on the cost function, for seven trials, normalized from zero to one for comparison. Translational or rotational offsets from the most accurate coregistration result in an increase in the summed ROI BP_{ND} cost function.

It is rarely the case that the PET and MRI images are misaligned by only a single translation or rotation. The results from simultaneously changing translation and rotation parameters (Figure 5) show that increasing angular perturbations can also affect the gradient of the cost function. For this test, the norm of the offset values was calculated as $\sqrt{x^2 + y^2 + z^2}$, where (x,y,z) were either the translation or rotation offsets (in mm or degrees, respectively) along each axis. The trend (left side of Figure 5) is an increasing cost function value as the coregistration becomes less accurate. The two plotted lines show, however, that as the rotational misalignment increases, the gradient of the cost function decreases and the residual error to the fitted line increases. In other words, the larger the initial rotational offset, the more difficult it is to overcome the translation error. To further visualize this point, residuals to the fitted trend lines were plotted for increasing rotational perturbations (right side of Figure 5). The increasing residuals confirm that a search over six-parameter space (three rotations, three translations) is less straightforward when the rotational error is high.

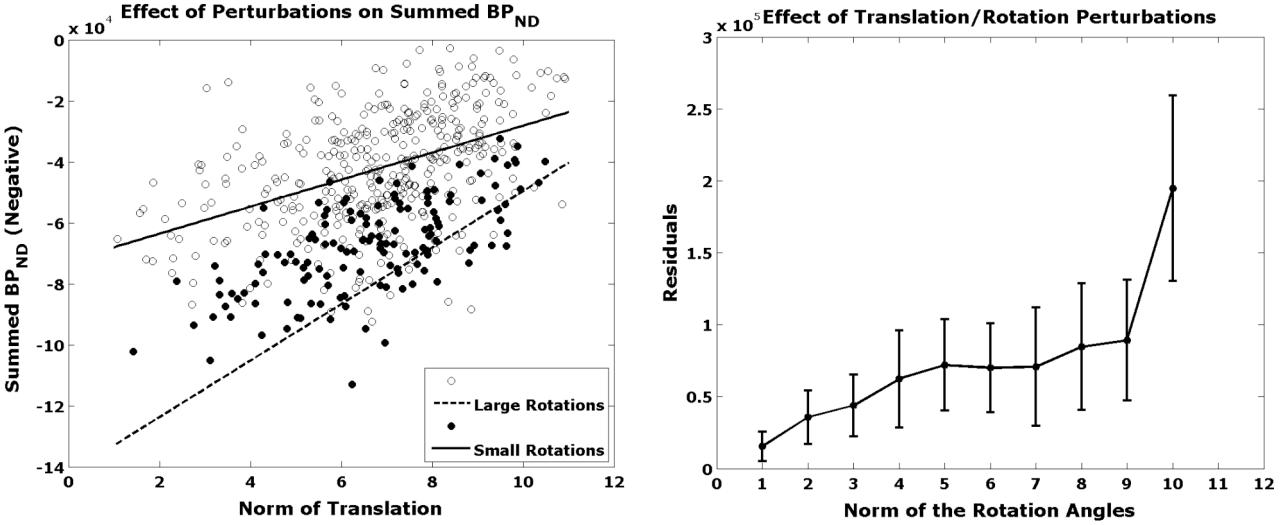

Fig. 5. **Left:** Simultaneous rotation and translation offsets. The closed and open circles are negative summed ROI BP_{ND} values from angular rotations with a norm equal to less than one or greater than ten degrees, respectively. The fitted lines show the trend of increasing cost function value as the translation perturbation increases. As the rotation perturbation increases, residuals to the fitted line increase and the slope decreases. **Right:** As the norm of the angular perturbation increases, so does the sum of the residuals.

3.3 Application to Test-Retest Study

Since the subject, ligand and imaging procedure are held fixed in a test/retest paradigm, these studies provide an excellent basis to evaluate the reproducibility of image processing procedures. Ideally, the same binding should be observed in the test and retest evaluations. This is not the case for a variety of reasons, such as image noise, motion artifact, image processing inaccuracies, or even ligand binding differences. Assuming intrasubject binding differences are minimal, the optimal image processing technique is the one that produces the most repeatable intrasubject binding measurements. Using the BP_{ND}-based coregistration technique produced repeatable binding estimates in the cauate, putamen and ventral striatum of the test/retest subjects. The mean intrasubject percent different (PD) in these regions was lower when the the BP_{ND}-based coregistration technique was used than when the best coregistration was defined as the one that produced the lowest cost function (of the eight coregistrations). See Figure 6 for the PD of each subject evaluated.

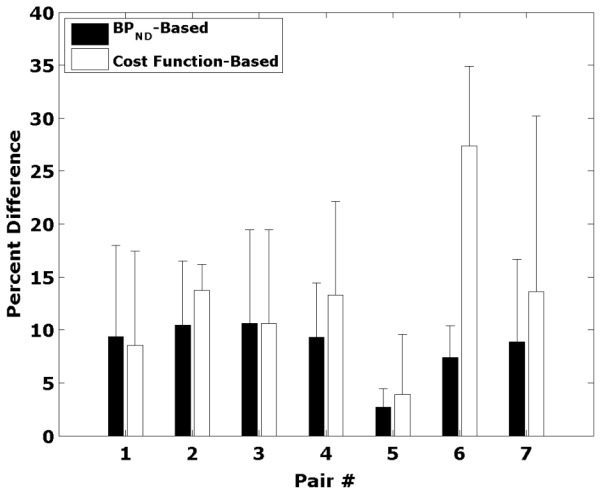

Fig. 6. Comparison of coregistration results. The BP_{ND} values of three high binding regions of interest were considered – dorsal caudate, dorsal putamen and ventral striatum. Bars represent the mean percent difference (PD) between test and retest BP_{ND} measurements, for each test-retest pair. Error bars indicate standard deviation of the PD in the three regions. Eight coregistration methods were performed for each PET image. The optimum coregistration was chosen as either the one that maximized BP_{ND} values (black bars) or produced the smallest cost based on a mutual information cost function (white bars). The chart shows that for six out of the seven test-retest pairs, choosing a coregistration based on BP_{ND} values produces a lower PD measurement.

4. CONCLUSIONS

The purpose of this study was to determine if V_T or BP_{ND} values could be used to assess coregistration accuracy. This is an important problem because PET/MRI misregistrations can cause significant differences in the calculation of both binding potential and volume of distribution, affecting decisions about a ligand's usefulness, receptor density, or dosimetry studies. Though coregistration accuracy can be assessed by eye, this task is laborious, difficult (as Figure 1 indicates) and prone to error (Table 2). However, we have demonstrated that the evaluation of maximal ROI BP_{ND} is a repeatable measurement with a level of accuracy at least as good as a manual selection. The novel contribution of this paper is the development and testing of this evaluation procedure for seven cases. Validation experiments indicate that, although it may be sensitive to large rotational misregistrations, summed BP_{ND} measurements within regions of interest perform well as an indicator of coregistration accuracy. This procedure would therefore be useful in place of (or perhaps in conjunction with) mutual information as a cost function. Future work on this project will extend the use of this evaluation procedure to serve as an alternative cost function for PET-MRI coregistration rather than as a coregistration selection method.

REFERENCES

[1] Knops, Z. F., Maintz, J. B. A., Viergever, M. A., and Pluim, J. P. W., "Normalized Mutual Information Based Registration Using k-means Clustering and Shading Correction," Medical Image Analysis 10(3), 432-439 (2006)

[2] Watson, C. C., Newport, D., Casey, M. E., "A single scatter simulation technique for scatter correction in 3D PET," [Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine], Aix-les-bains, France, 215-219 (1995)

[3] Mawlawi, O., Martinez, D., Slifstein, M., Broft, A., Chatterjee, R., Hwang, D. R., Huang, Y., Simpson, N., Ngo, K., Van Heertum, R., Laruelle, M., "Imaging human mesolimbic dopamine transmission with positron emission tomography. Part I: Accuracy and precision of D(2) receptor parameter measurements in ventral striatum," Journal of Cerebral Blood Flow & Metabolism 21, 1034–1057 (2001)

[4] Ogden, R. T., Ojha A., Erlandsson, K., Oquendo, M., Mann, J. J., and Parsey, R. V., "*In vivo* Quantification of Serotonin Transporters using [^{11}C]DASB and Positron Emission Tomography in Humans: Modeling Considerations," Journal of Cerebral Blood Flow & Metabolism 27, 205-217 (2007)

[5] Gunn, R. N., Sargent, P. A., Bench, C. J., Rabiner, E. A., Osman, S., Pike, V. W., Hume, S. P., Grasby, P. M., Lammertsma, A. A., "Tracer kinetic modeling of the 5-HT1A receptor ligand [carbonyl-^{11}C]WAY-100635 for PET," NeuroImage 8, 426-440 (1998)

[6] Innis, R. B., Cunningham, V. J., Delforge J., Fujita, M., Gjedde, A., Gunn, R. N, Holden J., Houle, S., Huang, S. C., Ichise, M., Iida, H., et al., "Consensus Nomenclature for In Vivo Imaging of Reversibly Binding Ligands," Journal of Cerebral Blood Flow & Metabolism 27, 1533-1539 (2007)

[7] Logan, J., "Graphical Analysis of PET Data Applied to Reversible and Irreversible Tracers," Nuclear Medicine & Biology 27, 661-670 (2000)