Assessment of the performance of SPM analysis in PET neuroactivation studies: A Monte Carlo investigation

Frédéric Schoenahl, Marie-Louise Montandon, Daniel O Slosman, Habib Zaidi

Abstract-- **Statistical parametric mapping (SPM) is now considered the gold standard both in clinical and research neuroimaging investigations. We investigate the feasibility of Monte Carlo simulations-based assessment of the sensitivity and specificity of SPM analysis in Positron Emission Tomography (PET) neuroactivation studies with respect to design parameters such as study size, activation/lesion localisation, intensity and size. Fifteen data sets were generated for different foci activation levels, and localizations using five spheres located within the digital Hoffman 3D brain phantom with diameters corresponding to 6 and 8 mm, close to the resolution limit of the ECAT ART PET scanner (CTI/Siemens). The five foci were defined in the *thalamus* (T), *putamen* (P), *cingulate gyrus* (CG), *right frontal cortex* (RFC) and the *right supramarginal gyrus* (RSG) covering different structures of particular interest in cognitive neuroscience. In addition, we carried out a coordinate search of the resulting maximum cluster, which was found to be close to 10 mm around the true value on the Talairach atlas.**

I. Introduction

STATISTICAL Parametric Mapping (SPM) - (Wellcome Department of Cognitive Neurology, University College London, London, UK) - refers to the construction and assessment of spatially extended statistical process used to test hypotheses about neuroimaging data. It is an academic software toolkit for analysis of functional brain images provided by a range of modalities commonly used in medical imaging. This package is widely used both in research and clinical investigations, mainly for detection of activation foci using SPECT, PET, fMRI and recently EEG data. SPM is considered actually among the state-of-the-art packages for statistical analysis of neuroimaging data because it is very flexible, well documented and extensively tested.

As the package remains open to any statistical analysis, very few publications reported on the validation of SPM analysis for limited number of subjects and to assess its performance for detection of characterized lesions, essentially small activation foci with varying intensity. There is a single experimental study [1] and few simulation studies in both PET [2] and SPECT [3] that were carried out to validate the methodology. Our approach uses a modified Hoffman 3D brain phantom and realistic Monte Carlo simulations to generate brain volumes corresponding to baseline and activation studies. The image generation model remains close to the reality, and multiple statistics is mimicked by changing the seed of the random number generator.

This work was supported by the Swiss National Science Foundation under grant SNSF 3152-062008.

The authors are with the division of nuclear medicine, Geneva University Hospital, Geneva 1211, Switzerland (telephone: +41 22 372 7258, e-mail: habib.zaidi@hcuge.ch).

II. Materials and Methods

A. Monte Carlo simulation of 3D brain PET data

The *Eidolon* Monte Carlo simulation package was used in this work [4]. To run simulations with sufficient statistics, it was decided to make some approximations to reduce the computation time required for data sets generation.

Preliminary studies were intended to optimize the simulator computation time. This was done in three steps, first by eliminating interaction of photons with detector blocks feature, and activating a convolution step simulating the response function of the PET tomograph. The spatial resolution of the tomograph was approximated by its diameter, crystal width and detector-block to PMT coupling effect, positron range and annihilation photons acolinearity [5] resulting in an intrinsic resolution of 4.5 mm. This corresponds to the design of the ECAT ART scanner [6] when scanning ^{15}O-[H_2O] patients. The second step involved a reduction of the polar angle through variance reduction techniques, which limited the polar emission angle of the annihilation photons in the axial direction without affecting the resulting scatter fraction. The acquisition was simulated in 3D mode with a maximum ring difference corresponding to 17 rings and a span of 7. Scripting facilities written in *GNU Bash* were used to perform independent simulations on workstations, and update parameters automatically. It was therefore possible to simulate 3x15 projection data with 300 Mcounts each in approximately one week.

B. Phantom construction

The numerical Hoffman 3D brain phantom [7] was used to generate the data sets on 3 Sun UltraSparc workstations. This phantom has been developed to simulate the activity distribution found in the human brain in typical ^{18}F-[FDG] metabolism studies currently employed in PET. The

mathematical source distribution has specific activities of the grey matter, white matter and ventricles of 4:1:0, respectively. Basic image processing utilities using MedX software (Sensor Systems, Inc.) were applied to make its characteristics match normal ^{15}O-[H_2O] cerebral blood flow distributions [2].

The numerical brain phantom was resampled to obtain a regular voxel-grid of 2x2x2mm, so that the implementation of a sphere drawing algorithm, written in C language was easier. The software was coupled to a script which allowed to insert the spheres at constant position within the raw phantom (Fig. 1). With the help of an experienced neuropsychologist, e added to the phantom a compartment for cerebrospinal fluid and set compartment relative concentration ratios to 0:2:25:100, for background, cerebrospinal fluid, white matter and grey matter, respectively.

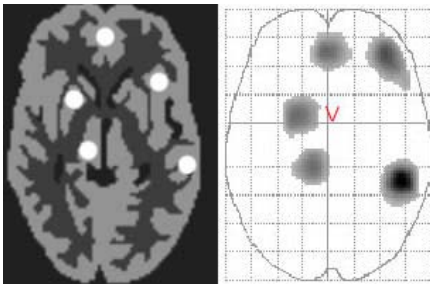

Fig. 1. Numerical brain phantom with artificial lesions inserted used as input for Monte Carlo simulations (left), and sample SPM map resulting from statistical image analysis (right).

To assess the sensitivity and specificity of SPM analysis in PET neuroactivation studies with respect to design parameters such as study size, activation/lesion localisation, intensity and size, fifteen data sets were generated for different foci activation levels, and localizations using five spheres located within the digital Hoffman 3D brain phantom with diameters corresponding to 6 and 8 mm, close to the resolution limit of the ECAT ART PET scanner. The five foci were defined in the *thalamus* (T), *putamen* (P), *cingulate gyrus* (CG), *right frontal cortex* (RFC) and the *right supramarginal gyrus* (RSG) covering different structures of particular interest in cognitive neuroscience.

C. *Image reconstruction and analysis*

Model-based scatter correction and calculated fit-ellipse based attenuation compensation and were applied to sinograms prior to reconstruction. The reconstruction of data sets followed the protocol commonly used in our division for brain scanning, namely the reprojection algorithm (3DRP) with cut-off frequency at 0.35 and standard ramp filter.

The reconstructed images were analyzed with SPM99, with a phantom-specific template built in house. We used the specifications of the template provided with SPM and coregistered manually our volume to the Talairach atlas. The main intention was to limit the influence of spatial normalization during SPM analysis as reported in [1]. This helps to avoid the distortion of the coregistered volume in the Z-direction due to missing slices in the Hoffman phantom at the neck level. Typical normalization results of simulated images are shown in Fig. 2 together with the template constructed for this purpose.

To study the effect of study design, the activity level of spheres, and the degrees of freedom (number of image pairs) used were varied, comparing sets from 3 to 15 images each with different activation levels. To keep the experimental set-up close to practical clinical and research investigations, we included a constant misregistration error to all activated sets, which remained inferior to the mean error ranges observed in SPM analysis.

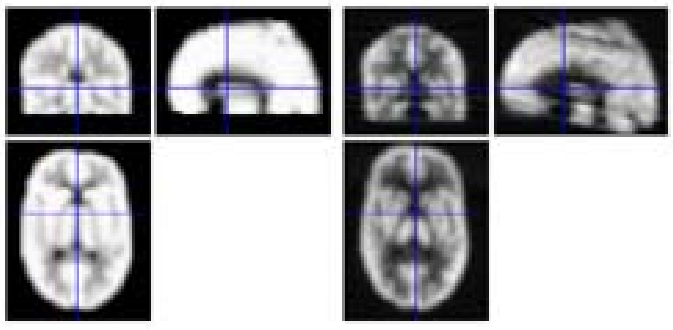

Fig. 2. Spatial normalization of simulated images (right) to the constructed template (left). It is shown here that the particular shape of the phantom requires a custom template for automatic normalisation. The "missing" proximal slices would induce significant transformation of the simulated data.

Spatial normalization was applied using the built-in template. The SPM analysis parameters are summarized in Table 1.

TABLE I
SUMMARY OF SPM ANALYSIS PARAMETERS.

Normalization	2x3x2 basis functions, 5 iterations (reduced influence of non-linear transformations, and faster computations)
Coregistration	The activated set manually misaligned, and its mean coregistered to the mean baseline set (PET > PET).
Smoothing	13 mm (~2xFWHM of simulated scanner)
Comparison	Simple *t*-test with two groups and one condition (50 ml/100 ml/min, 0.8 grey matter threshold)
Results	Analysed without and with threshold correction (*P*=0.05), extent threshold at 20 voxels.

III. RESULTS

We compiled the minimal activation thresholds required for detecting the different activation foci as function of the number of image pairs (Fig. 3). Sphere-clusters were considered as detected by the software only if they had higher probability than the first false positive found by the SPM software, if any. This rule was applied to mimic usual

applications where the user does not have prior information on the locations of activations and their extent.

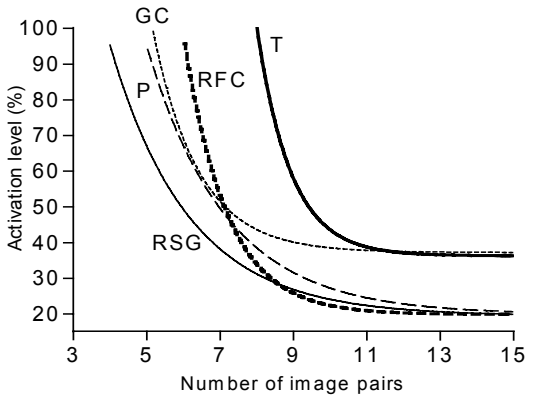

Fig. 3. Detection sensitivity for spheres of 8 mm at different locations within the brain phantom.

As expected, no cluster was detected with this design when analysing data sets corresponding to the smallest sphere with 6 mm diameter (resolution limit of the tomograph). It was observed that the correction of the height threshold could mask some lesions for small degrees of freedom (number of image pairs). A 6 mm sphere with 15 image pairs could only be detected with an uncorrected threshold.

It was also found that the depth of spheres within brain structures had an impact on detection sensitivity. Indeed, spheres located in the peripheral structures were detected more efficiently. This could be explained by the rough spatial normalization on spheres that were far from the anterior commissure, which are most affected by rotational transformations and from the peripheral cortex where small errors in spatial registration can lead to large tissue differences. The case of sphere located as *RSG* which leads to the best detection sensitivity can be explained by the thin size of cortex at this level and the complex structure of tissue interface the lesion has been put in. The resulting spill-over effect may enhance the signal.

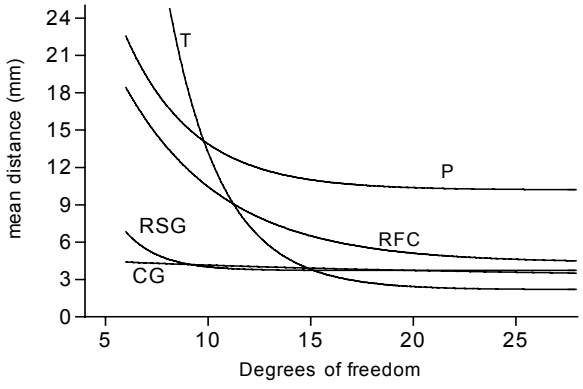

Fig. 4. "Coordinate search": mean distance from the true coordinate for spheres of 8 mm diameter. The maxima reach ~25 mm for the lowest degrees of freedom, the identity of Thalamus locus (max. 27 mm from original location) was deduced from its neighbour in Putamen, which is less sensitive to a decrease of the degrees of freedom.

As complementary analysis, we have compiled the coordinates of cluster maxima found by SPM for each level of activation. We calculated the Cartesian distance from the real sphere coordinates expected in the Talairach space for each sphere and level of activation. Results are summarised in Fig. 4 showing a mean difference comprised between 2 and 8 mm, which increases up to 25 mm for smallest usable degrees of freedom. The same logic was applied for comparison of each activation level with the expected location (Fig. 5).

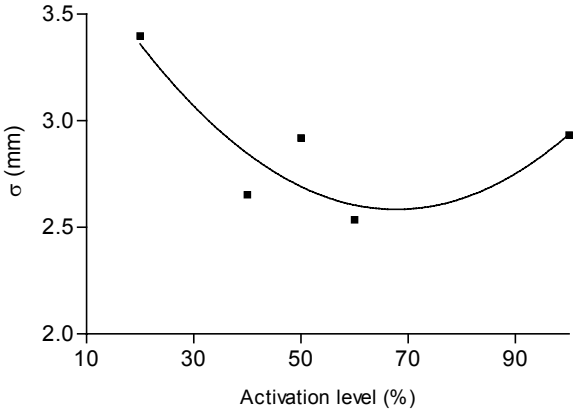

Fig. 5. Mean standard deviation (mm) from the expected location at different locations for the maximum number of image pairs (df = 18 and more). Results remain in the range 2.5 to 3.5 mm.

If the result could be influenced by the number of observations which varies for low degrees of freedom, we assume the statistical uncertainty of the result to be reduced by SPM's statistical correlation and the limited effect of normalization in our design (constant parameters for all activation sets, and adequate template). The fact that most intense areas were localized with good accuracy may suggest that the normalization was not influenced by the highest activation foci. However this effect could explain a relative low decrease of accuracy at highest activation levels (Fig. 5).

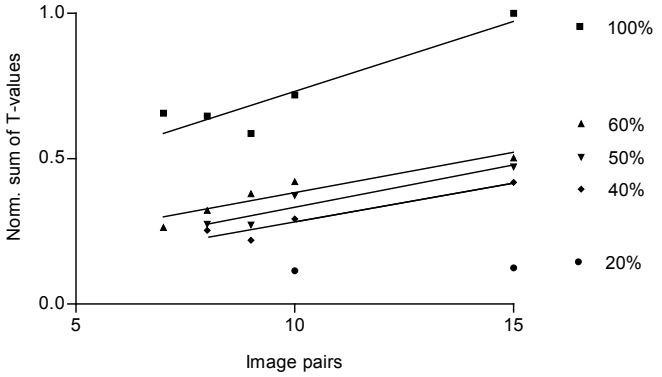

Fig. 6. Mean T-Statistic for all spheres with increasing degrees of freedom and activity level.

Another indicator was also calculated, namely the sum of T-statistics for each sphere. The results show large differences between e.g. *RSG* and *T* localizations (Fig. 6). Similar to the previous reasoning, the particular case of the lateral locus, which is located in the most heterogeneous medium, may indicate a spill-over effect. Fig. 7 shows the increase in T values for all individual localizations.

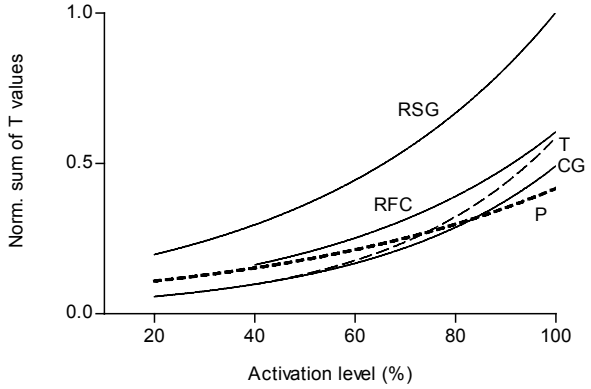

Fig. 7. Increase in T-values for individual localizations.

IV. Discussion And Conclusion

We have evaluated in this study the feasibility of Monte Carlo simulations-based validation of SPM software through detailed analysis of sensitivity and specificity of detection of simple spherical activation foci. The detection of 8 mm spheres for activation levels of 20% for low image pairs could be achieved on brain scans with relatively low statistics. The time required for simulations remains high to perform large datasets (28 degrees of freedom or more), but this performance could be greatly improved by using most recent multi-processor personal computers with 2 GHz and more. This is the reason why the simulator was ported on Linux platforms. The simulator could be validated against physical measurements performed on the PET scanner it has been parameterized for. The low statistic of data sets (~15 Mcounts per set) was close to the mean number of events collected during typical ^{15}O-[H_2O] dynamic brain scanning, We therefore expect the detection thresholds to be improved slightly with higher statistics.

The location of spheres clearly plays a role: if located close to a transformation center like the AC or lateral cortex, the probability is enhanced. Other locations (*T*, *RFC*) have lower detection statistics. Such an analysis could be extended to the assessment of the influence of activation level on normalization errors and the bias caused by high levels on coregistration, but this would require accurate correction for partial-volume/spill-over effects to reduce overestimation of intensities by SPM.

Results are in agreement with previously published studies using simulated and measured SPECT data [1, 3] considering the characteristic differences between PET and SPECT in terms of resolution and sensitivity performances. These preliminary data will be helpful in the design of future protocols in [^{15}O]-H_2O PET neuroactivation studies. More detailed analysis of the influence of different data correction techniques (attenuation and scatter, partial volume effects) and reconstruction methods (analytic, iterative) on the sensitivity and specificity of an SPM analysis is warranted. The effect of iterative reconstruction and model-based scatter correction on ^{18}F-[FDG] distribution of reconstructed 3D brain PET images was recently investigated using SPM analysis in healthy volunteers [8]. It was concluded that iterative reconstruction didn't results in significant changes while significant differences in ^{18}F-[FDG] distribution exist when images are reconstructed with and without explicit scatter correction for some cerebral areas. This needs to be acknowledged for adequate interpretation of 3D brain PET images after applying scatter correction.

It is foreseen that the methodology followed in this study will allow the assessment of the other parameters influencing the statistical analysis of neuroimaging data. In particular, the evaluation of other statistical comparison methods, such as the recently proposed Monte Carlo-based assessment of statistical inference, remains to be explored.

V. Acknowledgements

The authors would like to thank Dr K. Boulanouar and Prof. P. Celsis from INSERM U455, University of Toulouse III, France for their fruitful comments.

VI. References

[1] K. J. Van Laere, J. Versijpt, M. Koole, et al., "Experimental Performance Assessment of SPM for SPECT Neuroactivation Studies Using a Subresolution Sandwich Phantom Design.," Neuroimage, vol. 16, pp. 200-216, 2002.

[2] C. Davatzikos, H. H. Li, E. Herskovits, et al., "Accuracy and sensitivity of detection of activation foci in the brain via statistical parametric mapping: a study using a PET simulator.," Neuroimage, vol. 13, pp. 176-84, 2001.

[3] P. Lahorte, S. Vandenberghe, K. Van_Laere, et al., "Assessing the performance of SPM analyses of spect neuroactivation studies. Statistical Parametric Mapping.," Neuroimage, vol. 12, pp. 757-64, 2000.

[4] H. Zaidi, A. Herrmann Scheurer, and C. Morel, "An object-oriented Monte Carlo simulator for 3D positron tomographs.," Comput. Methods Programs Biomed., vol. 58, pp. 133-145, 1999.

[5] C. S. Levin and E. J. Hoffman, "Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution.," Phys Med Biol, vol. 44, pp. 781-799, 1999.

[6] D. L. Bailey, H. Young, P. M. Bloomfield, et al., "ECAT ART - a continuously rotating PET camera: performance characteristics, initial clinical studies, and installation considerations in a nuclear medicine department.," Eur J Nucl Med, vol. 24, pp. 6-15, 1997.

[7] E. J. Hoffman, P. D. Cutler, W. M. Digby, et al., "3-D phantom to simulate cerebral blood flow and metabolic images for PET.," IEEE Trans Nucl Sci, vol. 37, pp. 616-620, 1990.

[8] M-L. Montandon, D. O. Slosman and H. Zaidi, "Assessment of the impact of model-based scatter correction on ^{18}F-FDG 3D brain PET in healthy subjects using statistical parametric mapping.," Neuroimage (2003) *in press*