Statistical Power Maps for SPM Analysis of PET Scans

Juan D. Gispert[1], Santiago Reig[1], Javier Pascau[1], Monica Penedo[1], Juan J. Vaquero[1], Vicente Molina[2], Manuel Desco[1]

[1]Medicina y Cirugía Experimental. Hospital General Universitario Gregorio Marañón. Madrid, Spain., [2]Dept. of Psychiatry. Hospital Clínico Universitario. Salamanca, Spain.

INTRODUCTION

The statistical analysis of brain functional neuroimages (PET and fMRI) may result negative even though a true difference exists if the statistical test does not reach an adequate statistical power. When a negative result occurs, information about the statistical power achieved by the test is seldom provided. To a certain extent, this fact may explain the abundance of contradictory conclusions in the literature. A usual strategy to address this problem in other areas consists in reporting the results with confidence intervals (CI), which supply an estimation of the effect size while preserving the statistical significance information. However, there are no standardized procedures for displaying confidence intervals of statistic parametric maps. A solution may consist in reporting the statistical power of the test whenever a negative result is obtained. This approach has been previously proposed by Van Horn, et al. (1998) who developed a method for estimating the power of statistical parametric maps showing the probability of detecting as significant the effect size actually found.

This work presents an alternative method for reporting negative results in statistical parametric maps, consisting in estimating the maximum effect size that the test would not detect as significant with a certain probability.

MATERIAL AND METHODS

This maximum non-significant effect size is calculated as shown in Figure 1, in which t^{-1} denotes the inverse of Student's T cumulative distribution function, α the significance threshold of the statistical test, β the required statistical power, N_1 and N_2 the sample size of the two groups respectively and $_{s1}$ and $_{s2}$ its variances. We have applied this procedure to a statistical parametric map of a two-group comparison: ^{18}F-FDG PET images of 17 schizophrenic patients on classical neuroleptics were compared to those acquired six moths after switching their treatment to olanzapine.

RESULTS

The statistical parametric map revealed no significant differences ($\alpha = 0.001$) between the pre- and post-olanzapine groups. Figure 2 shows the estimation of the maximum non-significant effect size ($\alpha = 0.001$, $\beta = 0.80$) overlaid on a reference MR image.

CONCLUSIONS

If no statistical power analysis is performed when obtaining a negative result, the only correct inference that can be made is that the p-values are non-significant and the experiment is not conclusive. In these cases, concluding that there is no between-group difference is clearly erroneous, yet a common mistake. Our approach to assess statistical power differs from the previously reported method of Van Horn et al. who proposed to use the probability that the between-group differences would had been detected by the test. We believe that our method provides a more useful way to report statistical power since effect sizes can be directly compared between different studies thus facilitating the performance of meta-analyses.

$$d_{maxN.S.} = t_{\alpha} \cdot s_{pooled}$$

where

$$t_{\alpha} = t^{-1}(1-\alpha/2, N_1+N_2-2) + t^{-1}(1-\beta/2, N_1+N_2-2)$$

and

$$s_{pooled} = \sqrt{\frac{[(N_1-1)\cdot s_1 + (N_2-1)\cdot s_2] + \left[\frac{1}{N_1}+\frac{1}{N_2}\right]}{N_1+N_2}}$$

Figure 1.

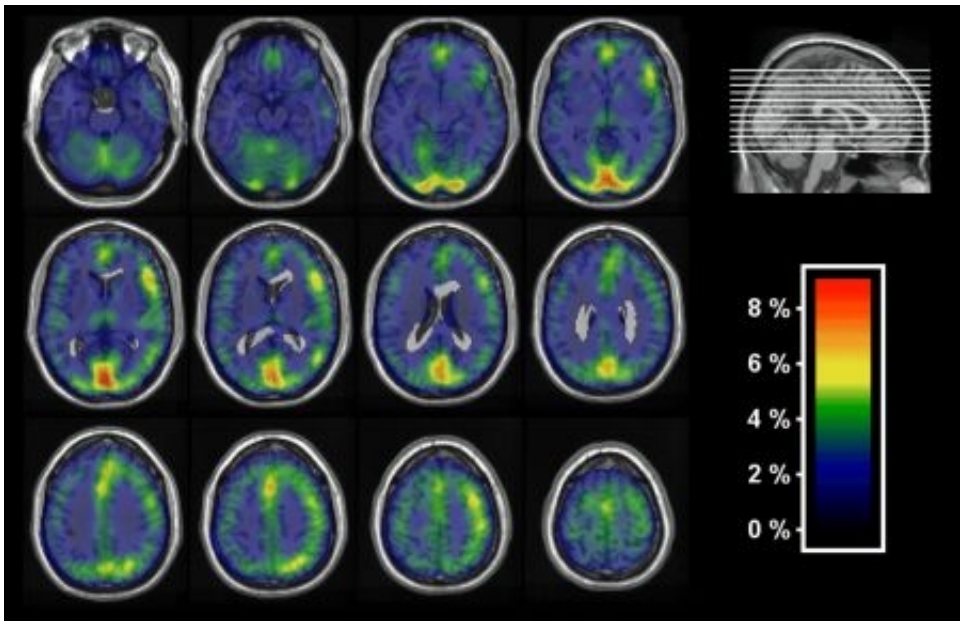

Figure 2: Maximum effect size not reaching statistical significance (α = 0.001; β = 0.80). The statistical power map is overlaid on a MR image for a better localization of the brain regions. This image allows us to discard (with a 80% probability) that changes above 9% of the global metabolism would occur anywhere between the two groups. Brain regions with lower statistical power are shown in red (visual cortex, anterior cingulated and left insula). In the rest of the cerebral cortex, between-group differences above 2% of the global metabolism could be discarded (blue regions).