How many subjects should be included in a well-powered cross-sectional cortical thickness analysis?

H. R. Pardoe[1], D. F. Abbott[1], and G. D. Jackson[1,2]
[1]Brain Research Institute, Florey Neuroscience Institutes, Melbourne, Victoria, Australia, [2]Department of Medicine, University of Melbourne, Melbourne, Victoria, Australia

Introduction Inter-subject registration of cortical thickness maps derived from whole brain T1-weighted MRI is a widely-used method for quantitative, objective analysis of focal neuroanatomical differences between subject groups. For the researcher designing a prospective study, a primary concern is to ensure that the study is adequately powered. Power analysis allows us to calculate the number of subjects required to minimise the probability of making a type II error (false negative). In this study we present the results of vertex-wise power analysis to calculate the number of subjects required to detect cortical thickness changes of a specified size, using standard methods for cortical thickness mapping and inter-subject registration. The number of required subjects is mapped back onto the target template, allowing us to investigate the spatial variability of the minimum number of subjects required for an adequately powered cross-sectional cortical thickness study across the cortical surface. Previous research [1] has provided sample-size estimates, however these analyses did not investigate the spatial distribution of minimum number of subjects over the cortical surface.

Material and methods 0.9 mm isotropic T1-weighted 3D MPRAGE MRI (Siemens TIM Trio) were acquired in 98 neurologically normal controls (53 female, age 29.1 ± 9.7 years). Cortical thickness mapping and inter-subject registration was carried out using the standard Freesurfer 5.0 processing stream [2]. Individual cortical thickness maps were registered to the supplied fsaverage template and smoothed using a surface-based smoothing filter with a spatial extent of 5,10,15,20 and 25 mm. The general linear model was used to estimate and correct for the effects of age and sex on vertex-wise cortical thickness estimates. T-test based power analysis was used to calculate the minimum number of subjects required to adequately control for the likelihood of a type II error. Standard deviation was estimated from the vertex-wise cortical thickness estimates. The following settings were used: power = 0.8, significance level = 0.05, two-sample T-test. Power analyses were conducted for hypothetical effect sizes of 0.125 mm, 0.25 mm, 0.375 mm, 0.5 mm, 0.625 mm, 0.75 mm, 0.875 mm, and 1 mm. One- and two-sided tests were used. One-sided tests are appropriate for studies in which the direction of change is already known (e.g. cortical thinning), whereas two-sided tests would be appropriate for exploratory studies without prior hypotheses regarding the direction of cortical differences. The calculated minimum sample size was mapped back onto the registration target to provide a whole-brain map of the minimum number of subjects required to reliably detect the given effect size. The PALS-B12 lobar atlas provided with the Freesurfer 5.0 distribution was used to estimate the number of subjects required to provide adequate power on a per-lobe basis. The relationship between the number of subjects in each group and effect size, and the effect of smoothing filter on the number of subjects required in each group for a given effect size was investigated.

Results The minimum number of subjects required for a well-powered cross-sectional cortical thickness analysis exhibits considerable heterogeneity over the surface of the cortex (Fig 1). Regions such as the anterior temporal lobe, insula and supra-marginal gyrus require considerably more subjects than other brain regions to achieve the same level of power. Using the PALS-B12 lobar atlas allowed us to estimate the number of subjects required for a well-powered study on a per-lobe basis. This analysis indicates that, in order to cover 95% of each lobe, the frontal, parietal and occipital lobes require approximately 30 subjects, whereas the temporal lobe requires approximately fifty subjects to detect a 0.25 mm thickness difference (Fig 2). The limbic lobe would require up to 234 subjects for an adequately powered study to detect the same effect size. In order to reliably detect a cortical thickness change of 0.25 mm over 95% of the entire cortical surface, around 60 subjects are required in each group with surface based smoothing of 10mm FWHM (Fig 3). For a thickness difference of 1 mm, less than ten subjects are required in each group. The spatial extent of the smoothing filter has a strong effect on the number of subjects required for a well-powered analysis, with the number of subjects required to detect 0.25 mm ranging from greater than 160 when no smoothing is applied to around 20 subjects when a large smoothing filter of 25 mm is applied (Fig 4). The size of the applied smoothing filter sensitizes the analysis to detecting cortical thickness differences of the same spatial scale, by the matched filter theorem. Therefore the values plotted in Figure 4 are applicable when the spatial extent of the thickness difference is equal to or greater than the spatial extent of the specified smoothing filter.

Discussion & Conclusions We have presented a method for calculating how many subjects should be included in a cross-sectional cortical thickness study in order to reliably detect a given effect size. Based on the results from a standard clinical scanner, the study has shown that the minimum number of subjects varies considerably over the cortical surface. If the researcher is interested in regions such as the anterior and medial surfaces of the temporal lobe, insula, supra-marginal gyrus and some other cortical regions, considerably more subjects will be required than the precentral gyrus or inferior parietal lobe, for example. By providing an explicit measure of how many subjects should be included in a study, applications of the results of this study include: (i) prospective study planning, (ii) quantifying how improvements in MRI acquisition and hardware can improve the sensitivity of brain imaging studies, (iii) meta-analysis of studies, and (iv) comparison of different cortical thickness mapping methodologies.

References [1] Han et al (2006) Neuroimage 32 pp. 180 - 194 [2] Fischl et al (2000) PNAS 97(20) pp. 11050 - 11055

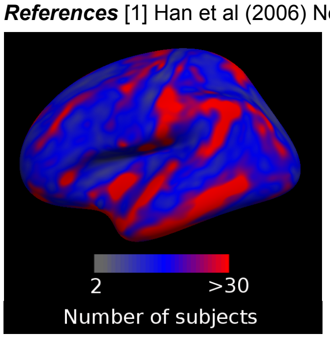

Fig 1. Left lateral inflated surface view of the vertex-wise estimates of the number of subjects required in each group to detect a difference of 0.25 mm after 10 mm smoothing (alpha = 0.05, two-sided test).

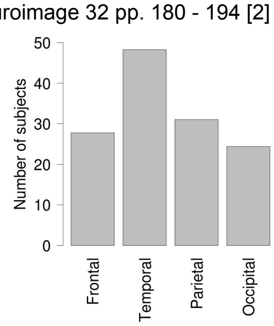

Fig 2. Number of subjects per group required for adequate power to detect a change of 0.25 mm over 95% of each major lobe after 10mm smoothing (two-sided test). The limbic lobe required 234 subjects and was omitted from the plot.

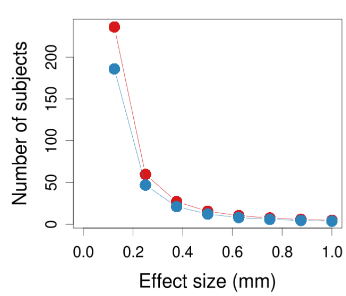

Fig 3. Number of subjects required to detect given effect size over 95% of the cortical surface. Two-sided (red) and one-sided (blue) results are shown. 10 mm surface-based smoothing kernel.

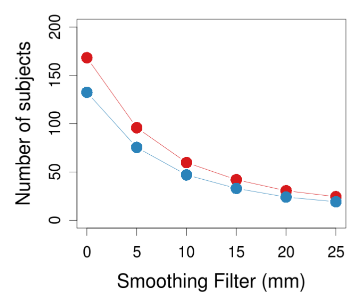

Fig 4. The number of subjects required in each group to detect a thickness difference of 0.25 mm following application of a smoothing filter with the specified spatial extent. Two-sided (red) and one-sided (blue) results are shown.