Available online at www.sciencedirect.com
ScienceDirect
Magnetic Resonance Imaging 26 (2008) 490–503
MAGNETIC RESONANCE IMAGING

Effects of spatial smoothing on fMRI group inferences

Michal Mikl[a,b,*], Radek Mareček[b], Petr Hluštík[c,d], Martina Pavlicová[e], Aleš Drastich[a], Pavel Chlebus[b], Milan Brázdil[b], Petr Krupa[f]

[a]Department of Biomedical Engineering, FEEC, Brno University of Technology, Koleni 4, 612 00 Brno, Czech Republic
[b]First Department of Neurology, St. Anne's University Hospital and Masaryk University, Pekarska 53, 656 91 Brno, Czech Republic
[c]Department of Neurology, School of Medicine, Palacky University, I.P. Pavlova 6, 775 20, Olomouc, Czech Republic
[d]Department of Radiology, School of Medicine, Palacky University, I.P. Pavlova 6, 775 20, Olomouc, Czech Republic
[e]Department of Biostatistics, Mailman School of Public Health, Columbia University, 722 West 168th Street, New York, NY 10032, USA
[f]Diagnostic Imagining Clinic, St. Anne's University Hospital and Masaryk University, Pekarska 53, 656 91 Brno, Czech Republic

Received 10 April 2007; accepted 20 August 2007

Abstract

The analysis of functional magnetic resonance imaging (fMRI) data involves multiple stages of data pre-processing before the activation can be statistically detected. Spatial smoothing is a very common pre-processing step in the analysis of functional brain imaging data. This study presents a broad perspective on the influence of spatial smoothing on fMRI group activation results. The data obtained from 20 volunteers during a visual oddball task were used for this study. Spatial smoothing using an isotropic gaussian filter kernel with full width at half maximum (FWHM) sizes 2 to 30 mm with a step of 2 mm was applied in two levels — smoothing of fMRI data and/or smoothing of single-subject contrast files prior to general linear model random-effects group analysis generating statistical parametric maps. Five regions of interest were defined, and several parameters (coordinates of nearest local maxima, *t* value, corrected threshold, effect size, residual values, etc.) were evaluated to examine the effects of spatial smoothing. The optimal filter size for group analysis is discussed according to various criteria. For our experiment, the optimal FWHM is about 8 mm. We can conclude that for robust experiments and an adequate number of subjects in the study, the optimal FWHM for single-subject inference is similar to that for group inference (about 8 mm, according to spatial resolution). For less robust experiments and fewer subjects in the study, a higher FWHM would be optimal for group inference than for single-subject inferences.

Keywords: fMRI; Spatial smoothing; Group inferences; SPM

1. Introduction

Functional magnetic resonance imaging (fMRI) is a noninvasive imaging method, which is used for the localization of active brain areas. The analysis of fMRI data involves multiple stages of data pre-processing before the activation can be statistically detected. These steps include correction of movement (after an estimate of movement parameters), spatial transformations into standard anatomical space, various kinds of spatial and temporal filtering or signal normalization, etc. Within each step, specific values must be chosen for each processing parameter. The ideal choice of the parameters depends on the type of fMRI paradigm, on the type of inference (single-subject, multi-subject), on the spatial scale (resolution of images) required for inferences and on the particular compromise between sensitivity and specificity of the analysis. Spatial smoothing is a very common pre-processing step in the analysis of functional brain imaging data. Smoothing is most often implemented as a convolution of the imaging data with a gaussian smoothing kernel described by a parameter of full width at half maximum (FWHM). This step is an important as well as a controversial operation with both advantages and disadvantages. Usually, some amount of smoothing is applied prior to estimate of image registration parameters because it allows better registration [1]. Smoothing of fMRI data prior to the statistical analysis can increase the signal-to-noise ratio

* Corresponding author. First Department of Neurology, St. Anne's University Hospital, Brno 656 91, Czech Republic. Tel.: +420 543 182 685; fax: +420 543 182 624.
E-mail address: michal.mikl@fnusa.cz (M. Mikl).

0730-725X/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.mri.2007.08.006

(SNR) and increase sensitivity to signals of specific shapes and sizes depending on filter design [1–9]. Smoothing is also useful in reducing resampling-related artifacts after image registration [10]. Note also that there is a constraint on the lower limit of smoothing that can be used because statistical inference in statistical parametric maps (SPMs) generally depends on the theory of gaussian fields and implicitly assumes that the data are good lattice representations of a smooth gaussian field [2,3,11,12]. This only holds when the voxel size is appreciably smaller than the smoothness. Worsley and Friston [2] suggest that the effective FWHM should be at least twice the size of the voxel. Spatial smoothing is also useful for suppressing the influence of functional and anatomical variability within and across individual subjects. Voxel-based methods for the analysis of functional neuroimaging data rely on the assumption that, after spatial transformation (realignment and normalization), all voxels are not only in the same anatomical reference space but that activations over subjects are expressed in the same location. This assumption is not valid in most cases: this is the reason for using smoothing. An undesirable side effect of smoothing is a partial-voluming artifact along the edges of the brain, where brain voxels become smoothed with no-brain voxels [1]. This results in a dark rim, which might be mistaken for hypoactivity. Maisog and Chmielowska [1] introduced a method for correcting for this effect. Further undesirable effects of smoothing are decreased effective spatial resolution, blurring and/or shifting of activations and merging of adjacent peaks of activation. Some recent studies analyzed the influence of spatial smoothing on fMRI brain activations results [6,7,13–15]. In these studies, most often, the influence of smoothing on several subjects in a single-subject (first-level) approach was presented. In contrast, the influence of smoothing on group results was analyzed by White et al. [7]. In most of the studies, gaussian kernels were used for spatial smoothing. Various parameters were assessed in the resulting activation maps: sensitivity and specificity; comparison with high-resolution unsmoothed data; spatial extent and merging of two distinct regions; and replicability of fine scale motor localizations.

In an extension of previous studies, this study presents a comprehensive view of the influence of spatial smoothing on fMRI group activation results. Some new aspects were studied. We recorded coordinates of local maxima in predefined regions of interest and their shifts with spatial smoothing; t value at local maxima; and other variables describing the fMRI results. Dependences of these characteristics and coordinates of local maxima on filter size (parameter FWHM) were studied and compared to theoretical or previous similar results.

2. Theory

The following section reviews some theoretical results of the effect of signal width, anatomical variability and filter width on the detectability of the signal (for more details, see Ref. [4]). In a typical fMRI study, data are measured from n subjects. Spatial smoothing using a gaussian filter kernel is applied to each image to obtain smoothness with an effective FWHM of w in each dimension. Consider a simple model for fMRI activation with volumetric data as presented in Ref. [4]. Denote the three-dimensional (3D) gaussian function with standard deviation σ by

$$\Phi(\sigma) = (2\pi\sigma^2)^{-3/2}\exp[-||x||^2/(2\sigma^2)], \tag{1}$$

where $w = \sigma\sqrt{8\log_e 2}$ is the FWHM and x is a vector of coordinates. The convolution of two gaussians is still gaussian with a resulting FWHM of $w = \sqrt{w_1^2 + w_2^2}$, where w_1 and w_2 are FWHMs of original gaussians. We consider several assumptions for this model: there is just one peak contained in the signal. The peak is modeled as the gaussian function $h(2\pi\sigma_S^2)^{3/2}\ \Phi(\sigma_S)$ where h is the signal height and $w_s = \sigma_s\sqrt{8\log_e 2}$ is the signal FWHM. Stationary white noise ε is used to model a noise component added to the signal. Additional smoothing (image reconstruction) has the effect of convolution of the signal plus noise with a gaussian point spread function $\Phi(\sigma_R)$ with FWHM of $w_R = \sigma_R\sqrt{8\log_e 2}$. Location of the resulting signal in group data is affected by several sources of intersubject variability: anatomical and functional variability, and registration error. In our paper, we will use the term *intersubject variability* for the total effect of all three sources of variability. That variability perturbs the location of activation of each subject about the common (expected) location. These perturbations are modeled as a 3D gaussian random variable with standard deviation σ_A. The resulting image, averaged over subjects (a large number n of subjects), can be written as

$$\begin{aligned}\text{img} &= [\text{signal}*\text{inter-subject variability}+ \\ &\quad \text{whitenoise}]*\text{reconstruction} \\ &= \lfloor h(2\pi\sigma_S^2)^{3/2}\Phi(\sigma_S)*\Phi(\sigma_A) + \varepsilon/\sqrt{n}]*\Phi(\sigma_R)\end{aligned} \tag{2}$$

where * denotes convolution. The maximum of SNR can be obtained when the filter width matches the perturbed signal width (matched filter theorem), that is when

$$\sigma_R = \sqrt{\sigma_S^2 + \sigma_A^2}. \tag{3}$$

If we consider real data with several activation peaks with various heights, widths and distances between neighboring peaks, and we use various widths of smoothing kernel, then there are four theoretically possible alternative events according to changing selected filter width [16]: annihilation — an activation blob disappears; merge — two activation blobs merge into one; split — one activation blob splits into two; and creation — a new activation blob appears. The first two events often become arguments against the use of smoothing. Merging of two activation blobs will occur if the distance between peaks is less than twice the FWHM of the smoothing kernel.

Another theoretical issue is the impact of smoothing on common thresholding techniques. It is known that using massively univariate voxel-wise statistics requires correction for multiple testing. There are some common techniques to control family-wise error (FWE) [17] and some of them depend on spatial smoothness. SPM2 uses a combination of Bonferroni correction and the random field theory (RFT) to control FWE. The Bonferroni method is based on correction of the *P* value according to the number of multiple tests. For functional MRI studies, this correction seems to be very strict. The RFT method accounts for dependence in the data. Nichols and Hayasaka [17] have demonstrated that for a large observed *z* value, the corrected *P* value depends on the measure of search region and the measure of roughness (the inverse of smoothness). An increase in smoothness will cause a decrease in the corrected *P* value and an increase in significance. Since the roughness parameter is difficult to interpret, Worsley proposed a reparametrization in terms of the convolution of a white noise field into a random field with smoothness that matches the data. See more details in Refs. [4,17]. Worsley defined *resel* as a spatial resolution element with dimensions $FWHM_x \times FWHM_y \times FWHM_z$. Subsequently, we can combine search volume and resel size and instead we can use the search volume measured in resels. The RFT results then depend only on this single quantity, the resel volume. Note that RFT results are only valid for data that are smooth enough to approximate continuous random fields.

3. Methods

3.1. Subjects

Twenty healthy right-handed subjects participated in the study (7 males and 13 females). The mean age was 23± 3.9 years (with a range from 20 to 34 years; median, 22). The subjects were volunteers from the professional or academic sector, with no history of neurological or psychiatric disease. Informed consent was obtained from all subjects after all of the procedures were fully explained, and the study received the approval of the local ethics committee. Czech was the first language of all subjects. Handedness was determined according to subject preference for writing and drawing, which was determined by subject report and by direct observation.

3.2. Task

A visual oddball task was performed. In this task, a train of equally spaced visual stimuli is presented to the subjects. There are two types of stimuli: standard stimuli and target stimuli. The standard events occur more frequently than the targets. The subjects are instructed to count the target stimuli in their head and report the total number at the end of the experiment. In the present study, the standard visual stimulus=frequent event (93.7% of trials) was an image consisting of the string of white characters ‘OOOOO’ on a dark background, while the target image (6.25% of trials) was the string of white characters ‘XXXXX’ [18]. Visual stimuli were delivered via data projector onto the projecting screen and were seen by the subjects through a mirror that was mounted on the MRI scanner’s radiofrequency head coil. A total of 1024 images were shown to the subjects (64 targets and 960 standards) in four experimental runs of 256. The interstimulus interval was fixed at 1600 ms. The duration of stimuli exposure was constant at 500 ms. During the remaining time (~1100 ms), the screen was dark. The targets were distributed randomly among the four runs and 1024 trials, but it was ensured that there were at least eight frequent events between every pair of target events.

3.3. Image acquisition

Imaging was performed on a 1.5-T Siemens Symphony scanner equipped with the Numaris 4 System (MRease). Functional images were acquired using a gradient echo, echoplanar imaging sequence: TR (scan repeat time)=1600 ms, time to echo (TE)=45 ms, field of view (FOV)=250 mm, flip angle=90°, matrix size 64×64, slice thickness=6 mm, 15 transversal slices per scan. Each functional study consisted of a four runs, and each of the time course series consisted of 256 scans (total 1024 scans per subject). Following functional measurements, high-resolution anatomical T1-weighted images were acquired using a 3D sequence that served as a matrix for the functional imaging (160 sagittal slices, resolution 256×256 resampled to 512×512, slice thickness=1.17 mm, TR=1700 ms, TE=3.96 ms, FOV=246 mm, flip angle=15°).

3.4. fMRI data analysis

The SPM2 program (Wellcome Department of Imaging Neuroscience, London, UK) was used for processing the data. The following pre-processing was applied to each subject’s fMRI data: realignment to correct for any motion artifacts; normalization into the standard stereotactic space (MNI); spatial smoothing using a gaussian filter kernel with FWHM sizes 2 to 30 mm with a step of 2 mm, temporal filtering with a high-pass filter of 136 s and correction for serial correlations. The original imaging matrix generated from the above acquisition parameters was resampled to isotropic voxels with a size of 3 mm. For each subject and all sizes of FWHM, statistical parametric maps were computed to detect activation using a general linear model voxelwise analysis. Vectors of onsets of experimental conditions convolved with a kernel that approximates the hemodynamic response curve were used as regressors modeling specific effects of interest in the imaging data. A specific hypothesis (target>frequent) was tested with a *t* value [SPM (*t*)] at each voxel. Another objective of this study was to examine the effect of spatial smoothing applied to subject “results” (contrast files). Although the smoothing is a linear operator and generation of contrast maps using ordinary least squares

is a linear procedure, we were interested in verification of their permutation within SPM2 with real data. The rationale is that SPM2 uses restricted maximum likelihood estimation with autocorrelation estimates, which is not commutable with the smoothing operator. Thus, the secondary smoothing was applied to the contrast files entering into the group analysis. Subsequently, group results were analyzed using random-effect analysis (one-sample t test) for all of the combinations of primary (first-level) and secondary (second-level) smoothing.

3.5. Obtaining the characteristics of smoothed data

Five regions of interest (ROIs) were defined at the commonly observed locations [18–20] of the most significant activations during this kind of fMRI task (see Table 1). An FWHM of 6 mm was used as the initial value for the group analysis to generate coordinates of local maxima in these ROIs; a filter size of 6 mm was chosen as it is the typical value used in single-subject studies. Because of the large amount of group results (generated through combinations of primary and secondary smoothing), custom Matlab scripts were written to extract the required characteristics of interest. For each ROI, we were interested in the coordinates of the nearest local maxima; t value, contrast values (differences between target and frequent stimuli) and residual mean square values at these coordinates; and mean t value calculated from a spherical region centered on these coordinates (radius of 7.5 mm). For each group result, we were interested in the FWE corrected threshold, the false discovery rate (FDR) corrected threshold (both using a probability level of .05), the resel size, the resel count (the number of resels in the brain mask) and the voxel count (the number of voxels in the brain mask). The dynamic range of significant t statistic values was assessed by subtracting the FWE threshold from the t statistics at local maxima.

To study the possible resulting shifts of local maxima, absolute and relative euclidian distances between initial and actual coordinates of each local maximum were calculated. Absolute distances were always calculated to initial coordinates (both distances to coordinates of local maxima in unsmoothed data and distances to the initially selected coordinates from Table 1 within each ROI were calculated) while relative differences were calculated to the coordinates of local maxima obtained for the nearest lower smoothing.

Table 1
Description of ROIs

ROI	Initial MNI coordinates (x, y, z)	MNI coordinates of nearest local maxima for unsmoothed data	Intersubject variability (standard deviation/ FWHM)
IPS	48, −42, 48	45, −42, 48	3.03/7.14
ACC	9, 15, 39	6, 21, 36	3.15/7.40
SMA	3, 9, 54	0, 9, 54	2.57/6.05
IPL	57, −45, 21	57, −45, 18	2.98/7.00
ThR	15, −3, 15	15, 0, 15	2.27/5.34

The latter can be useful to find the filter widths generating significant shifts of peaks. Differences between maximum or mean statistics and FWE or FDR thresholds were calculated as well to examine their dependence on spatial smoothing. To separate the effect of intersubject variability from the effect of spatial smoothing, we calculated the measure of intersubject variability for each ROI. The distances between group coordinates of local maxima in specific ROI and single-subject coordinates of local maxima from unsmoothed data were calculated. Then, the intersubject variability was assessed as the standard deviation of these distances.

4. Results

4.1. Inspection of statistical parametric maps

We used maximum intensity projection (MIP) for a preliminary evaluation of thresholded statistical parametric maps (Fig. 1). If we used unsmoothed data, then there were no clearly distinct activation centers. When a typical or larger FWHM (>6 mm) was used to smooth the data, several distinct activation centers appeared. The t statistics were also more robust and the MIP images were less transparent. It was difficult to observe significant shifts of activation peaks. Moreover, differentiating between impact of first-level and second-level smoothing is difficult from MIP pictures. From this kind of inspection, it seems that the results are dependent only on the overall amount of smoothing.

4.2. Influence of spatial smoothing on t statistics

The graphs of the influence of spatial smoothing on t statistics at local maxima are illustrated in Fig. 2. Fig. 2A and C shows the influence of first-level smoothing (applied to MR data) on t values at local maxima and on mean t values from sphere region centered at the coordinates of local maxima, respectively. Fig. 2B and D shows the influence of second-level smoothing (applied to contrast files of single-subject results prior to group analysis) on t values and on mean t values, respectively. In Fig. 2A and B, there are distinct peaks for intraparietal sulcus (IPS), anterior cingulate cortex (ACC) and supplementary motor area (SMA) with FWHM sizes of 8 to 10 mm, whereas for inferior parietal lobule (IPL) and thalamus R (ThR), the first distinct peaks appear with FWHM of 8 and 6 mm, respectively, and additional peaks appear with FWHM of 16, 22 and 24 mm. We can see converging of plotted t values for large filter sizes, but it does not necessarily mean merging of peaks in activation maps. To resolve this problem, we have to compare the coordinates of ROIs. There are no important differences between first-level and second-level smoothing. The graphs of mean t statistics around local maxima are smoother, and instead of peaks there are plateaus. Fig. 2E shows the dependency of t statistics on a combination of first-level and second-level smoothing while the FWHM for second-level smoothing is

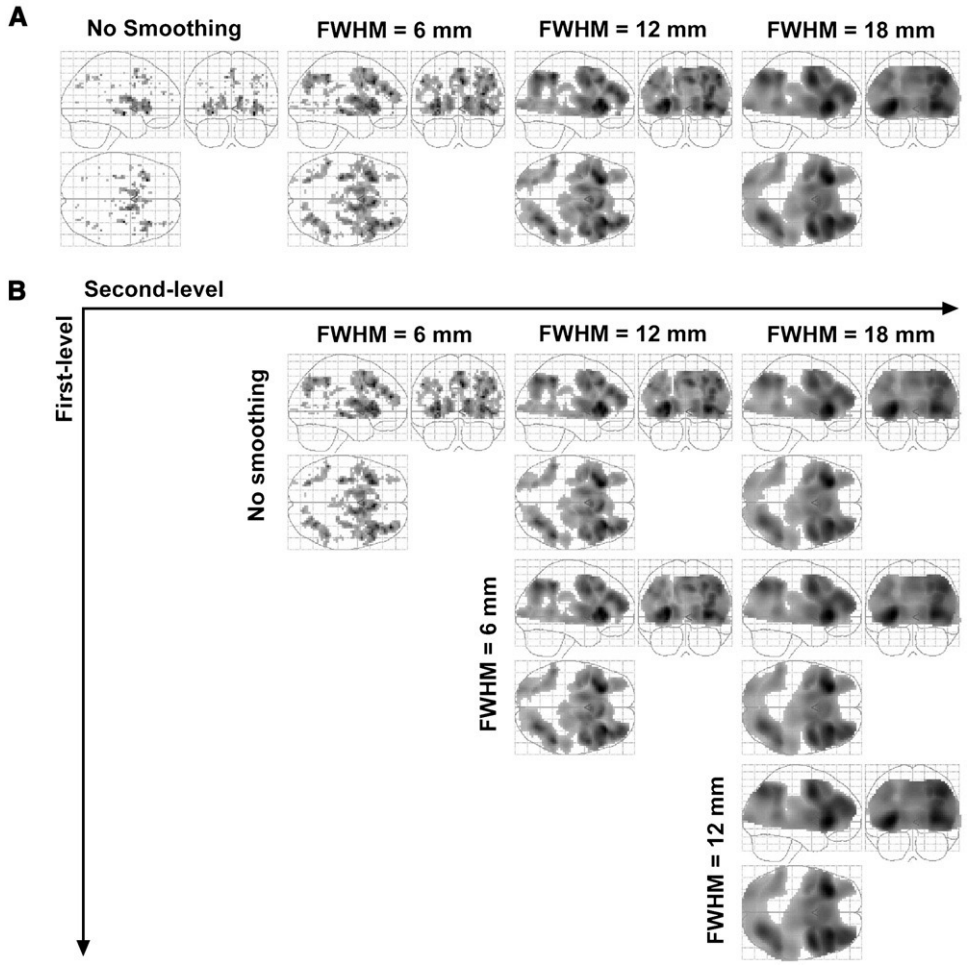

Fig. 1. Influence of spatial smoothing on MIPs of activation maps using a significance threshold of $P<.05$ corrected for multiple comparisons controlling FWE. (A) Only first-level smoothing (smoothing raw data) was used. (B) Combination of first-level and second-level (smoothing of contrast files entering into the second-level analysis) smoothing was used.

constant with a size of 30 mm. The last part of Fig. 2F shows the impact of spatial smoothing on the dynamic range of significant t statistic values. No significant difference against Fig. 2A was observed.

4.3. Influence of spatial smoothing on statistical thresholds and effective resolution of the data

Each graph of the FWE threshold as a function of FWHM (Fig. 3A and C) can be described as consisting of two sections. For the smaller FWHM, the statistical threshold is constant. For larger filter widths, the graph resembles an exponential decay. The reason for that is the combination within SPM2 of two methods for the calculation of FWE corrected threshold (Bonferroni correction and the RFT). For more details, see the Discussion section.

FDR corrected thresholds as a function of spatial smoothing are plotted in Fig. 3B and D. We can also see a small decrease in threshold with increased smoothing, which is not quite what was expected. In principle, the FDR method is not dependent on the spatial smoothness of the data. Thus, the reason for that is probably a change in calculating P values with various amounts of smoothing applied to the data.

As expected, we can observe the influence of spatial smoothing on effective resolution of the data (Fig. 4). There are slight differences between resel size for first-level smoothing and resel size for the same amount of second-level smoothing. A very interesting result is shown in Fig. 4E. We can see differences between voxel count (number of voxels) for first-level and second-level smoothing. The mechanism of this will be discussed in the Discussion section.

4.4. Influence of spatial smoothing on localization of local maxima

The effect of spatial smoothing on the localization of activation peaks is shown in Fig. 5. There are two types of graph. Panels 5A and B show euclidian distances between coordinates of local maxima and initial (reference) coordinates for each ROI (see Table 1). Thus, the zero or minimum

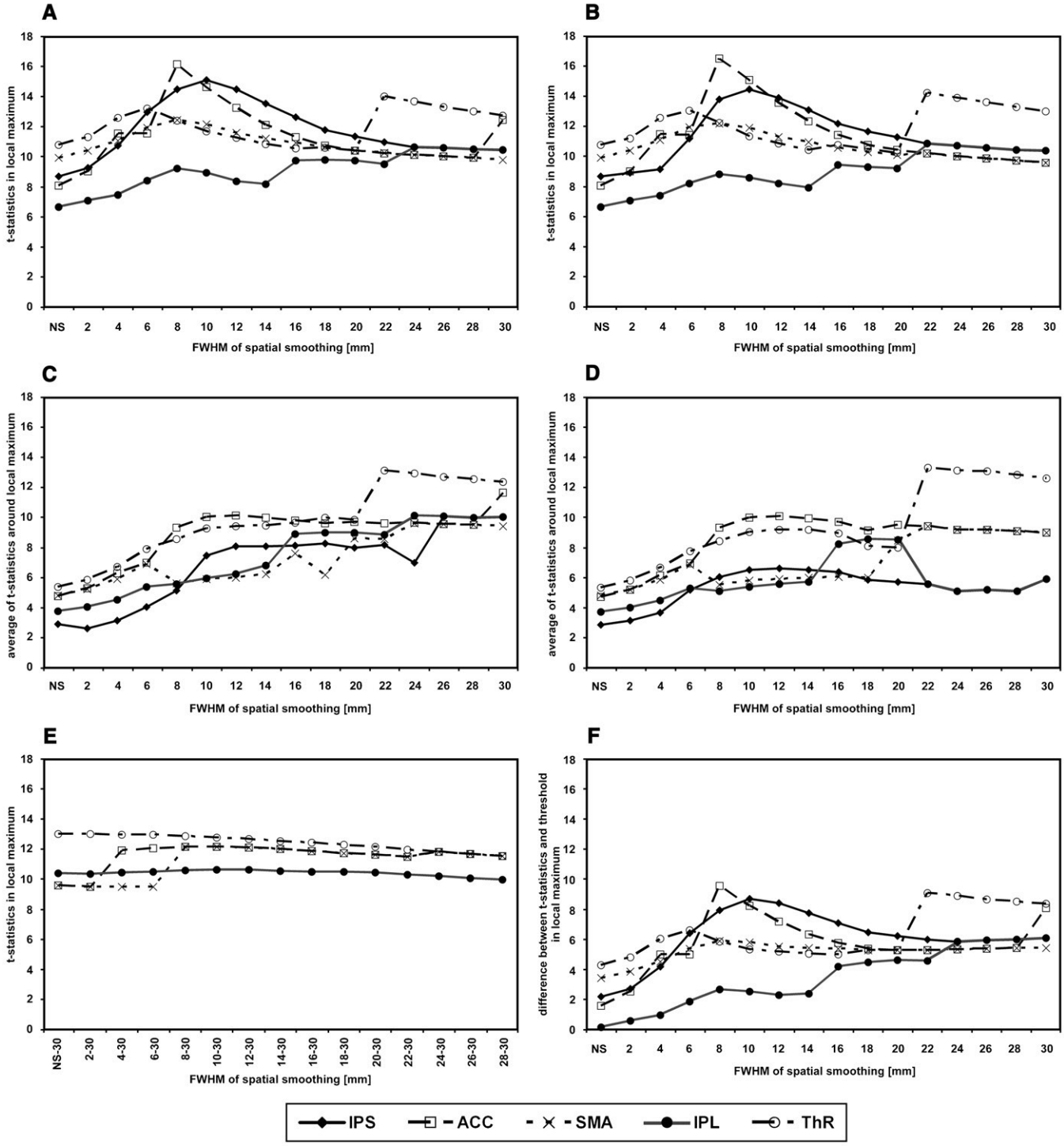

Fig. 2. Effect of smoothing on the *t* statistics. (A and B) The *t* values in the nearest local maximum. (A) Only first-level smoothing was applied. (B) Only second-level smoothing (smoothing of contrast files prior to second-level analysis) was applied. (C and D) Average *t* statistics around local maximum. (C) First-level smoothing only. (D) Second-level smoothing only. (E) Combination of first-level and second-level smoothing (the second-level FWHM is always 30 mm). The *t* statistics from local maximum are plotted. (F) Influence of first-level smoothing on the difference between *t* statistics in local maximum and FWE corrected threshold. It can be calculated as subtraction of values in Fig. 3A (FWE threshold) from values in A (*t* statistics in local maximum).

euclidian distances correspond to an FWHM of 6 mm. A different approach was chosen for Panels 5C and D, where the initial point is the same as the local maxima for unsmoothed data. The local maximum is very stable in the whole range of filter sizes for SMA, and it is relatively stable (shift up to 5 mm) up to an FWHM of 14 mm for a majority

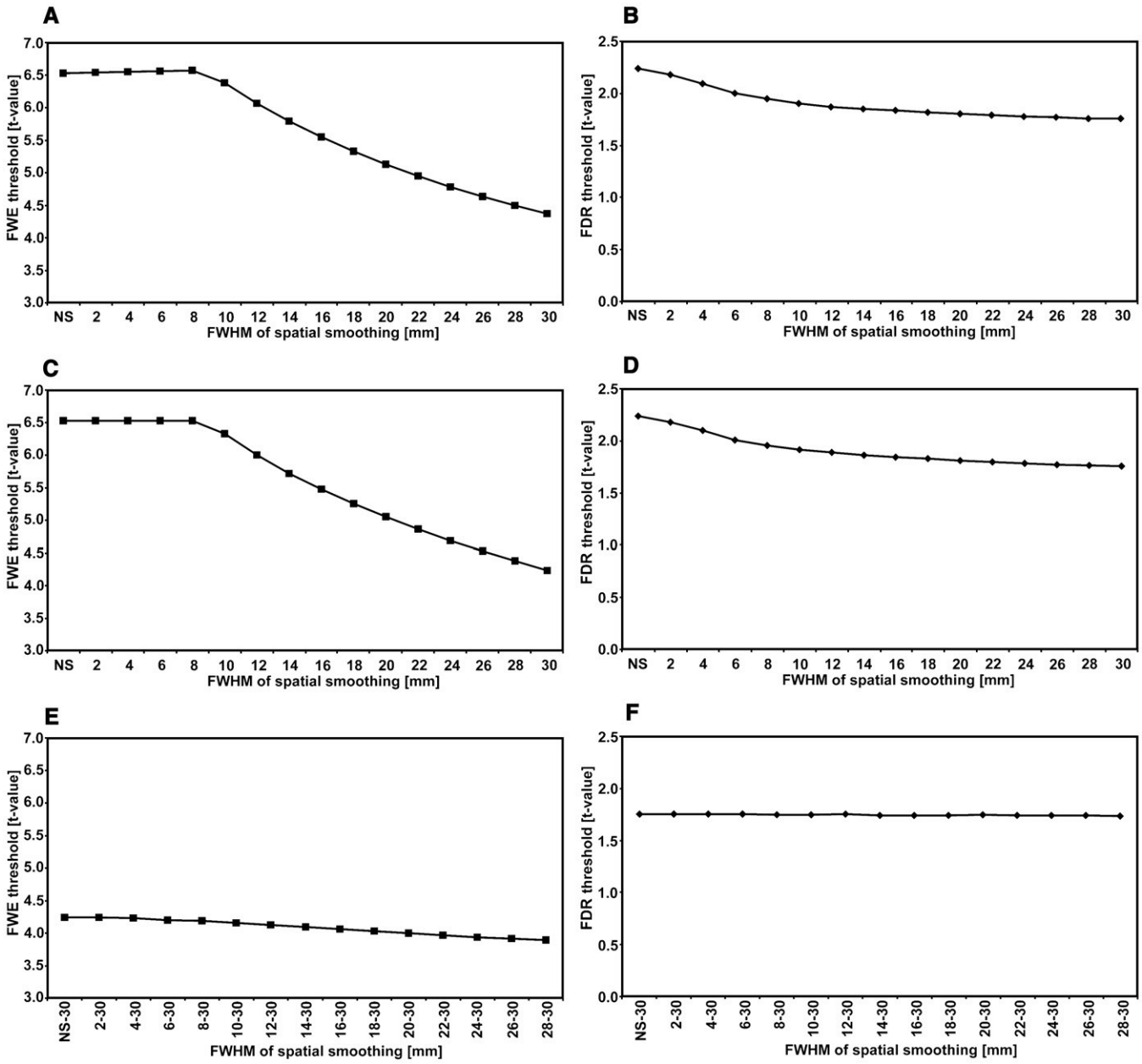

Fig. 3. Impact of spatial smoothing on whole-brain estimated statistical thresholds calculated by SPM2. (A and B) Only first-level smoothing was applied. (C and D) Only second-level smoothing was applied. (E and F) Combinations of first-level and second-level smoothing (the second-level FWHM is always 30 mm). (A, C and E) FWE threshold. (B, D and F) FDR threshold.

of ROIs except ACC where the shifts are up to 10 mm. The relative differences (Fig. 5F) show where (at which filter size) the distances change the most. We can see that for the majority of ROIs (except ACC), the significant shifts occur for an FWHM above 14 mm.

For that large smoothing, one of the causes is often the merging of neighboring activation peaks. During first-level smoothing, merging of our ROIs occurs at an FWHM of 26 mm (merging of IPS and IPL as well as merging of ACC and SMA). During second-level smoothing, merging occurs at an FWHM of 22 mm (the same pairs of ROIs). The euclidian distance between IPS and IPL is about 32 mm for unsmoothed data and about 29 mm for initial coordinates of ROIs (data smoothed with an FWHM of 6 mm). The distance between ACC and SMA is about 22 mm (unsmoothed data) and 17 mm (6 mm smoothing).

4.5. Influence of spatial smoothing on effect size (contrast values) and residual values

We consider the contrast between the effects of frequent and target stimuli as effect size. Fig. 6 shows that the effect size decreases almost linearly but residual values decrease

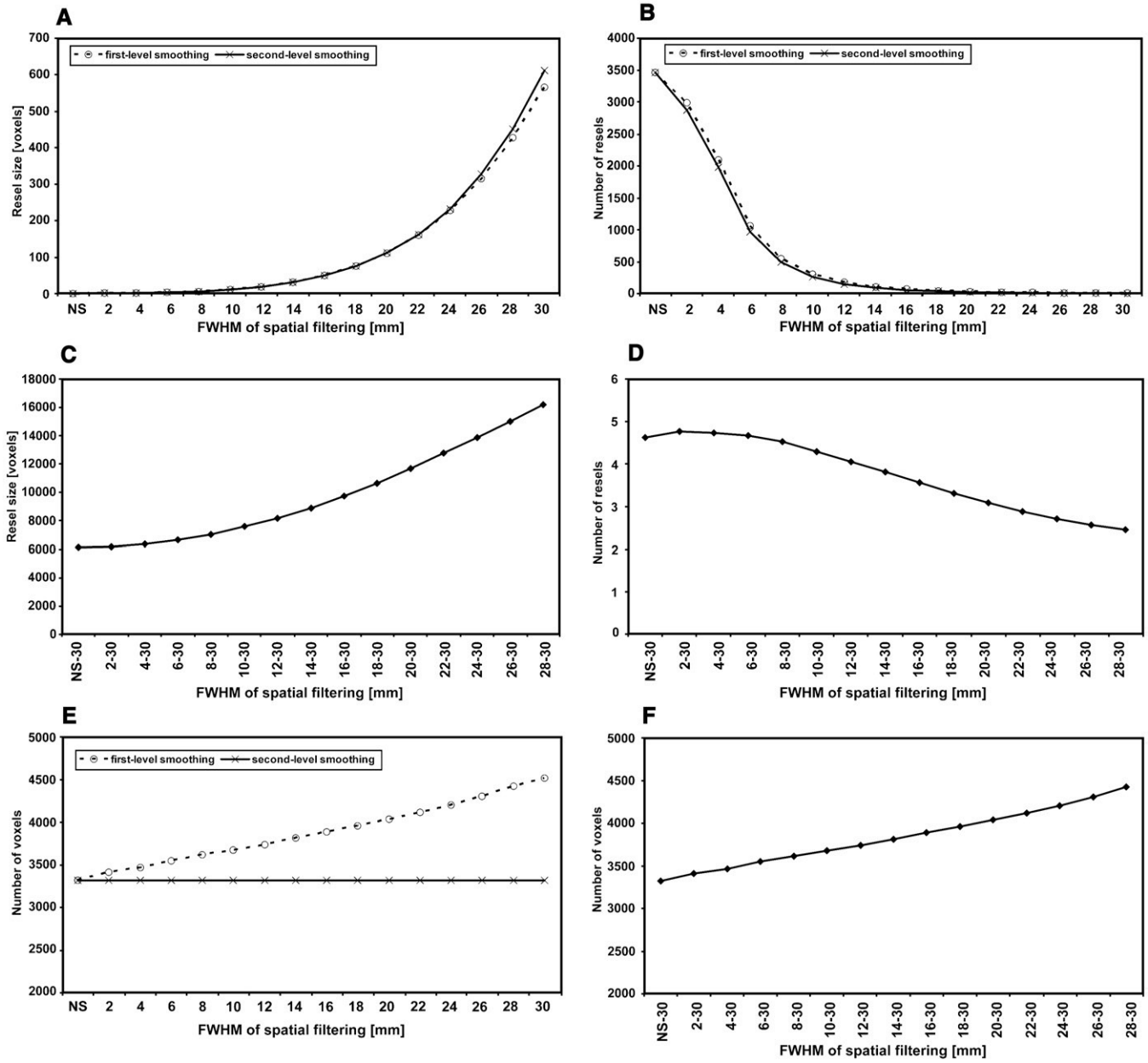

Fig. 4. Influence of spatial smoothing on the smoothness of fMRI results. Effects of smoothing on resel size, resel count and voxel count are plotted in graphs. (A, B and E) Either first-level or second-level smoothing was applied (see legend in graphs). (C, D and F) Combinations of first-level and second-level smoothing (the second-level FWHM is always 30 mm). (A and C) Influence of smoothing on resel size. (B and D) Influence of smoothing on number of resels. (E and F) Influence of smoothing on the number of voxels in the brain mask.

quickly with an approximately exponential decay. This explains the behavior of the *t* values seen in Fig. 2. We can say that the initial increase in *t* statistics is due to a rapid decrease in residual values with spatial smoothing. Note that the large values of effect size are due to linear combination of four sessions.

4.6. Comparison of intersubject variability with effect of spatial smoothing

For each ROI the intersubject variability was calculated from unsmoothed data using standard deviation of distances between group and single-subject local maxima. These standard deviations of intersubject variability were also converted to FWHM. The results are summarized for each ROI in Table 1.

Because the real data may violate theoretical assumptions, we provide three additional graphs of contrast and statistical values plotted along axes *x*, *y* and *z* presented in Fig. 7. These extractions of 3D data show the shape of the activation peak ("line profile") in the IPL region (initial coordinates are marked with dotted lines). We can see that the shape of peak is not gaussian, it is not very symmetrical and isotropic and

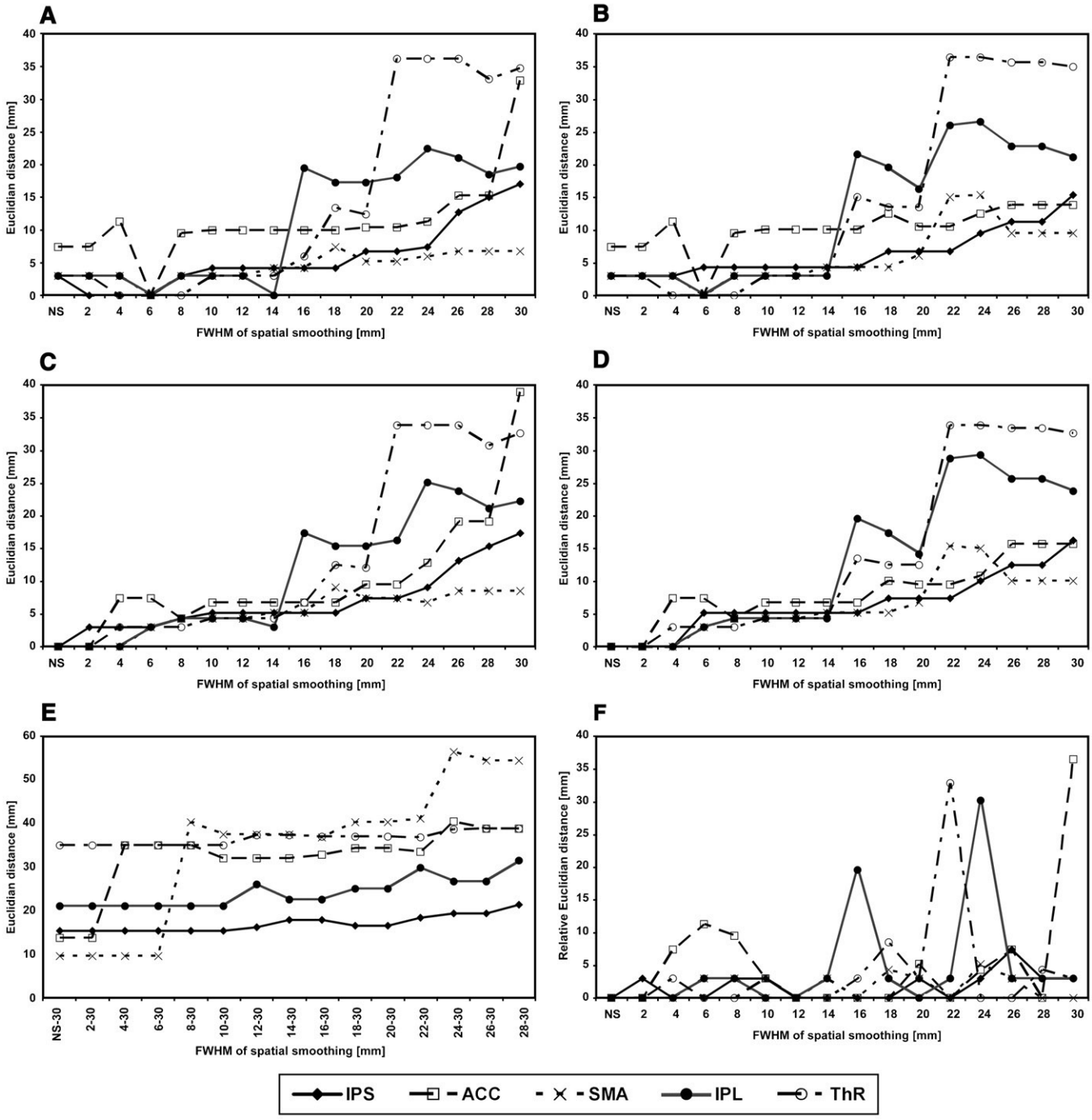

Fig. 5. Influence of smoothing on euclidian distances between the nearest local maximum of *t* statistics and its initial coordinates within each ROI. (A, C and F) First-level smoothing only. (B and D) Second-level smoothing only. (E) Combinations of first-level and second-level smoothing (the second-level FWHM is always 30 mm). (A, B and E) Coordinates from Table 1 (typical coordinates for selected ROIs) were selected as initial coordinates for calculating distances. (C, D and F) Corresponding distances when initial coordinates of each local maximum are derived from rough (unsmoothed) data. In part (F), relative differences are plotted (subtraction of euclidian distance for actual FWHM and for nearest lower FWHM). This graph highlights filter sizes introducing the largest spatial shift of each local maximum.

there is another smaller peak near the primary one. Thus, both the calculation of intersubject variability and the possible assessment of peak size (e.g., using FWHM) are not very accurate to deal with in the theoretical model. We have to use these characteristics only as an approximation.

5. Discussion

In this article, we present an empirical study of the influence of spatial smoothing on group SPM results. We considered several issues. The primary issue is to find the

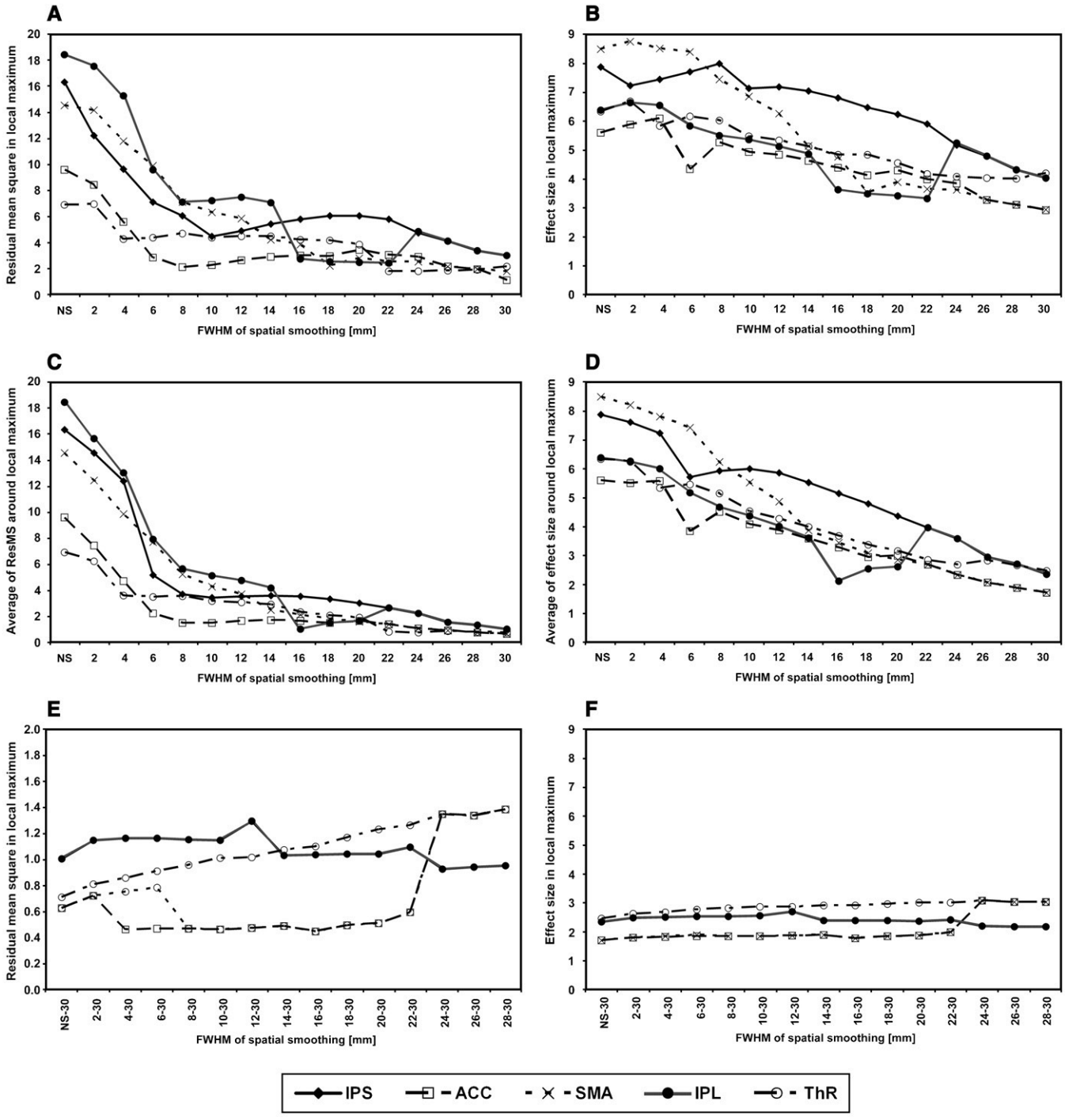

Fig. 6. Influence of smoothing on the residual mean square values (Graphs A, C and E) and the effect size (contrast values) (B, D and F). (A and B) Only first-level smoothing was applied. (C and D) Only second-level smoothing was applied. (E and F) Combinations of first-level and second-level smoothing (the second-level FWHM is always 30 mm). Graphs C and D show average characteristics around local maximum of *t* statistics. The others show characteristics from local maximum of *t* statistics. Note that the effect size is obtained from linear contrast summed across four experimental sessions.

rules for a rational choice of the optimal amount of spatial smoothing. The specific values we derived can be applied to similar fMRI tasks, but the more general rules can be used for a wide range of fMRI experiments. We also addressed the influence of spatial smoothing on various parameters or characteristics of the resulting statistical parametric maps and compared empirical data to known theory. Hence, many parameters and their dependence on spatial smoothing were studied and compared to theoretical or previous similar results.

Inspection of activation maps in MIP yields only basic general impressions. As the filter width increases, the spatial

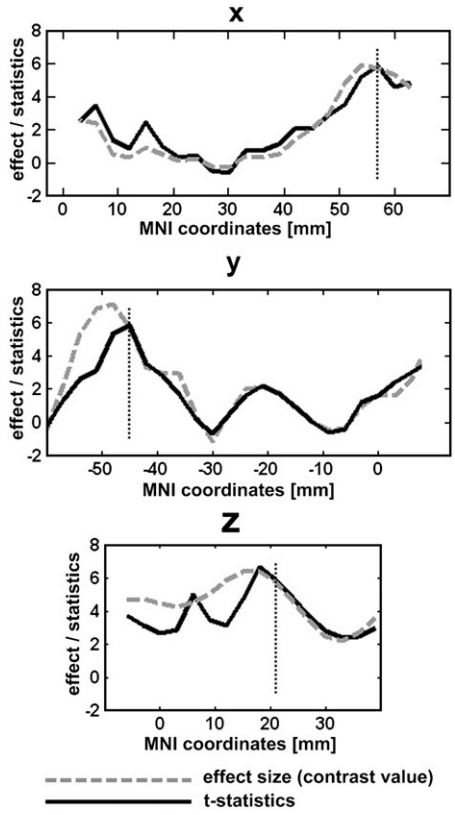

Fig. 7. Graphs of contrast (effect size) and statistical values plotted along axes *x*, *y* and *z*. These profiles extracted from the unsmoothed 3D data show the real shape of the activation peak in the right IPL region (MNI coordinates 57, −45, 21). These coordinates are marked with dotted lines. The *t* statistics is plotted with solid line and the effect size is plotted with dashed line. Note that the effect size is obtained from linear contrast summed across four sessions.

extent of the activations becomes larger and the statistics become more robust. But one cannot exactly judge whether there are any significant spatial shifts of activation centers. Also, the degree of significance of the noticeable changes is uncertain as long as we do not know the "true" underlying functional/anatomical representations.

We consider the influence of spatial smoothing on *t* statistics to be one of the most important characteristics investigated in this study. One possible optimization goal is to maximize sensitivity. To this end, we can search for maximum *t* values (or any other statistics used for activation detection) within predefined regions of interest. Such a method has been used in previous studies [6,13] but only at the first level of inference (single subjects). At the second level of inference, we must consider intersubject variability [4,7,21,22] in addition to single-subject results. The intersubject variability can be divided to anatomical and functional variability, and registration error. From the group-results point of view, these sources cause the same localization error among subjects and can be considered as an overall effect of intersubject variability. Higher sensitivity is then anticipated at larger filter sizes. White et al. [7] used a Hanning filter with an FWHM of 0 to 18 mm to study the effects of spatial smoothing on group fMRI data analysis. They suggest that an FWHM of 8 to 10 mm is appropriate to suppress the intersubject variability in the sensorimotor cortex but that a smaller filter size (an FWHM of about 4 mm) is more suitable in the thalamic and cerebellar areas. This result is quite consistent with our present data (we found maximal sensitivity in the thalamus with an FWHM of 6 mm and in other ROIs with an FWHM of 8–10 mm). In a preliminary study [23] with a different task and less robust data (only 71 scans per person, 20 subjects in the study, epoch design), the optimal FWHM for group inference was about 12 mm.

For the purpose of finding maximum sensitivity, we can interpret these results in two ways. First, considering there is no intersubject variability (as at the single-subject level of inference), we can conclude that we found the smoothing filter width that matches the true spatial extent of activation. Second, considering that the spatial extent of activation peaks is substantially smaller than intersubject variability, we can say that we found the best filter size to suppress intersubject variability. In real data, differentiation between these two effects is difficult, and it is maybe not necessary to separate them. We tried to calculate intersubject variability for each ROI. In all ROIs, the FWHMs of intersubject variability are smaller than FWHMs that were detected empirically as the filter size for the best sensitivity. This is expected, but if we consider an ROI size of the order of several millimeters, then slightly higher differences between intersubject variability measure and the most sensible filter size are expected [see Eq. (3)]. These can be due to violated assumptions that are necessary to support the theoretical model framework. Nonetheless, we can conclude that intersubject variability is suppressed at filter sizes smaller than those associated with maximal sensitivity. Filter size then depends on sample size (number of subjects in the study) and the robustness and the reproducibility of results for single subjects. Our experimental results seem to be quite robust and stable.

To control false positives, a correction for multiple testing is commonly used. It is important that the correlations of adjacent voxels in statistical processes are taken into account when solving the multiple comparison problem [24]. There are two main concepts: FWE and FDR, of which FWE can be realized in several ways. FWE implemented in SPM2 software uses two alternative approaches: the Bonferroni correction and the RFT [17], and the smaller of the two results is then used as the corrected threshold. The Bonferroni correction uses the number of voxels in the brain as the number of independent tests, whereas the RFT uses the resel volume (count of independent elements in the image data) and hence the RFT corrected threshold is dependent on the smoothness of the data. For example, in Ref. [25], the smoothness of statistical maps generated from unsmoothed, null-hypothesis data was estimated to be less than a single voxel (FWHM≈0.90 pixels). For this reason,

we assessed the influence of spatial smoothing on FWE and FDR thresholds, on resel size and on the number of resels (which represents the effective resolution or smoothness of the data). Our results are consistent with simulations presented in Ref. [17]. In addition to that, we used these observed dependencies for comparison of the dynamic range of t statistic values (subtraction of the thresholds from the maximum or mean t values) with the behavior of the maximum or mean t values. Those results are not reported in full as they are very similar (overall shape and position of peaks) to the graphs presented in Fig. 2. If the smoothed data are thresholded using the corrected thresholds, then one can argue that the main reason for achieving robust statistics is a decrease in the corrected threshold; however, this is only a minor (albeit still important) effect. The major effect originates from the changes in effect size and residual values. Note that although FDR correction is supposed to be independent of the voxel count and spatial smoothness of the data, it is possible to see a gradual decrease in FDR threshold with increasing smoothing. This is probably due to a change in the calculation of P values with various amounts of smoothing applied to the data. Here, it is necessary to discuss differences between observed numbers of (brain) voxels for first-level and second-level smoothing. The volume of the brain is constant, but the differences are caused by creating a brain mask. During this process, the mask is created according to the histogram of measured data to separate brain tissue from its surroundings. This is useful to decrease the number of statistical tests. But the brain mask is only generated after first-level smoothing (prior statistics), hence averaging of edge-of-brain voxels with out-of-brain voxels increases the total number of brain voxels with an increasing amount of spatial smoothing. For that reason, the mask created prior to first-level statistics was used after second-level smoothing due to suppression of more blurring activations outside the brain. Thus, second-level smoothing did not affect the number of voxels.

The influence of spatial smoothing on effect size and residual values was also studied. Variability across subjects decreases more rapidly than the mean effect over the subjects in the FWHM range from 2 to 8 mm. This is the reason for the noted increase of t values. Only a small decrease of residual values can be observed in the FWHM range from 8 to 30 mm. This is the reason (note that the effect size still decreases) for the appearance of peaks in the plots of maximum t statistics. The dependence of effect size on FWHM is more linear and decreases for all filter sizes. From this point of view, it seems that the FWHM of about 8 mm is optimal for suppressing intersubject variability. Knowledge of graphs of mean effect size and residual mean square values could be useful in deciding whether the final extent of activation is due to low intersubject variability and large activated areas at the single-subject level or due to high intersubject variability and the small extent of activation spread across single subjects. Worsley et al. [4] presented simulations to show the effects of intersubject variability and spatial smoothing on the signal-to-noise ratio. They demonstrated that if the anatomical variability equals the signal width, then the signal-to-noise ratio drops by 0.41, which is equivalent to losing 65% of subjects.

Geissler et al. [15] presented a study of replicability of the motor hand center with nonsmoothed and smoothed data analysis. After they used smoothed data, they observed an increase in motor center aberrations (spatial shifts) between repeated measurements of about 100%. In our study, spatial locations of activation centers were observed as well, but a different approach was chosen. Coordinates of the nearest (from the initial coordinates) local maximum within each ROI were recorded. Subsequently, two different types of reference coordinates were defined: initial coordinates for each ROI and coordinates obtained at the nearest local maxima when unsmoothed data were analyzed (see Table 1). Both approaches are presented in Fig. 5. Our results are useful for answering questions such as, “What is the maximum filter size that can be used without having the shift of activation center reach a certain critical distance?” or “What is the filter size that leads to the maximum/minimum distance of the activation center from its reference coordinates?” High variability was found across our predefined ROIs. We have to note that all meaningful filter sizes cause shifts of local maxima. Therefore, we should not use smoothing or use very low FWHM for applications that need accurate localization. On the other hand, if we tolerate spatial shifts of several millimeters (e.g., up to 5 or 6 mm), we can use spatial smoothing with an FWHM up to 12 mm. This value is only valid for similar experiments and numbers of subjects, but we can conclude that similar spatial smoothing will produce an acceptable amount of spatial shifts for almost all studies that are designed with respect to detection power and sample size appropriate for group inference.

One typical negative effect was observed in our results — the merging of ROIs. With first-level smoothing, merging of IPS and IPL at an FWHM of 26 mm as well as merging of ACC and SMA was noted. With second-level smoothing, merging of the same pairs of ROIs was noted at an FWHM of 22 mm. Hence, such a filter size may be inappropriate for nearby activations, but it is beyond the sensible filter size. The other possible events (annihilation, split and creation of peaks) were not observed in results from our ROIs. One possible explanation is the choice of ROIs or robustness and power of such an experimental design.

There are no considerable differences between first-level and second-level smoothing. This suggests that the mechanism used by SPM2, restricted maximum likelihood estimation with autocorrelation estimates, although in principle noncommutable, does not cause significant divergence of results processed either way. This appears to be useful for methodological practice. The suggestion that it is not necessary to re-process all subjects at first-level analysis may be particularly useful. We must note that the classical approach of spatial smoothing with a fixed isotropic filter

kernel is still commonly used (probably in most fMRI studies), but several concepts were published more recently in an attempt to improve spatial smoothing. One concept (presented originally for functional PET studies) is called "searching scale space" [4]. The authors proposed a search over a range of filter widths to find four-dimensional local maxima in location and scale space. This has the added advantage of estimating the signal width as well as its location. However, we think this concept is more effective for increasing SNR (matching the activation width) than for suppressing intersubject variability. A similar concept is a computer vision tool referred to as the scale-space primal sketch [16]. A rather different approach was presented as "Anatomically Informed Basis Functions" [26] for single-subject studies and for multisubject studies [22]. They used various forms of prior anatomical knowledge (based on reconstructed gray matter surfaces) and assumptions about the location and spatial smoothness of the blood oxygenation level-dependent signal to specify sophisticated spatiotemporal models for an fMRI time series. An approach that should remove the effect of nonlinearity of spatial smoothing on voxel variances (e.g., smoothing over different types of tissue) called masked contrast images [27] was introduced as a complement for standard methods of statistical mapping. Friman et al. [8] introduced novel and fundamental improvements to fMRI data analysis called "Adaptive Analysis of fMRI Data". These improvements included adaptive spatial filtering based on the concept of spatial basis filters. The above-mentioned articles provide important improvements to spatial smoothing. However, the classical approach (fixed isotropic filter size) is still used and implemented in most of the software for fMRI data analysis. It is for this reason that we tried to extend current knowledge about the effects of spatial smoothing on fMRI group analysis.

6. Conclusion

The optimal filter size for the group analysis of functional MRI depends on various criteria and specific functional areas and experimental tasks. From the sensitivity point of view, the optimal FWHM is 8–10 mm. These results are consistent with other studies. From the smoothness-dependent threshold point of view, it seems to be theoretically advantageous to use a higher FWHM, but the advantages are not practically significant if a thresholding procedure is implemented similar as in SPM2. To use RFT, it is only necessary to apply minimal smoothing (twice the voxel size). From the intersubject variability point of view, the optimal FWHM seems to be about 6–8 mm. From the spatial shift point of view, acceptable FWHM can be about 12 mm if we tolerate spatial shifts of up to 6 mm; otherwise, we should not use smoothing if we do not tolerate any spatial shift of activation peaks. This article evaluated many parameters and their dependence on spatial smoothing and attempts to contribute to a better understanding of the effects of spatial smoothing on group analysis. We must note that exact values of optimal FWHM are valid only for our experiment and similar experiments, but we suggest a more general conclusion as follows: for robust experiments with a large enough sample size (number of subject above 16), the optimal FWHM is similar for both single-subject inference and for group inference and it is about 8 mm (according to the spatial resolution of the data). For less robust experiments with a smaller sample size, the optimal FWHM will be higher for group inference than for single-subject inference and it will probably be about 10–14 mm according to intersubject variability, robustness of the data and spatial resolution.

Acknowledgments

The study was supported by MŠMT ÈR Research Program no. MSM0021622404.

References

[1] Maisog JM, Chmielowska J. An efficient method for correcting the edge artifact due to smoothing. Hum Brain Mapp 1998;6(3):128–36.
[2] Worsley KJ, Friston KJ. Analysis of fMRI time-series revisited-again. Neuroimage 1995;2(3):173–81.
[3] Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SC, Frackowiak RS, et al. Analysis of fMRI time-series revisited. Neuroimage 1995;2(1):45–53.
[4] Worsley KJ, Marrett S, Neelin P, Evans AC. Searching scale space for activation in PET images. Hum Brain Mapp 1996;4(1):74–90.
[5] Poline JB, Worsley KJ, Evans AC, Friston KJ. Combining spatial extent and peak intensity to test for activations in functional imaging. Neuroimage 1997;5(2):83–96.
[6] Xiong JH, Gao JH, Lancaster JL, Fox PT. Assessment and optimization of functional MRI analyses. Hum Brain Mapp 1996;4(3):153–67.
[7] White T, O'Leary D, Magnotta V, Arndt S, Flaum M, Andreasen NC. Anatomic and functional variability: the effects of filter size in group fMRI data analysis. Neuroimage 2001;13(4):577–88.
[8] Friman O, Borga M, Lundberg P, Knutsson H. Adaptive analysis of fMRI data. Neuroimage 2003;19(3):837–45.
[9] Scouten A, Papademetris X, Constable RT. Spatial resolution, signal-to-noise ratio, and smoothing in multi-subject functional MRI studies. Neuroimage 2006;30(3):787–93.
[10] Maas LC, Renshaw PF. Post-registration spatial filtering to reduce noise in functional MRI data sets. Magn Reson Imaging 1999;17(9): 1371–82.
[11] Friston KJ, Holmes A, Poline JB, Price CJ, Frith CD. Detecting activations in PET and fMRI: levels of inference and power. Neuroimage 1996;4(3 Pt 1):223–35.
[12] Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. A unified statistical approach for determining significant signals in images of cerebral activation. Hum Brain Mapp 1996;4(1):58–73.
[13] Hopfinger JB, Buchel C, Holmes AP, Friston KJ. A study of analysis parameters that influence the sensitivity of event-related fMRI analyses. Neuroimage 2000;11(4):326–33.
[14] Fransson P, Merboldt KD, Petersson KM, Ingvar M, Frahm J. On the effects of spatial filtering-a comparative fMRI study of episodic memory encoding at high and low resolution. Neuroimage 2002;16(4): 977–84.
[15] Geissler A, Lanzenberger R, Barth M, Tahamtan AR, Milakara D, Gartus A, et al. Influence of fMRI smoothing procedures on replicability of fine scale motor localization. Neuroimage 2005; 24(2):323–31.

[16] Lindeberg T, Lidberg P, Roland PE. Analysis of brain activation patterns using a 3-D scale-space primal sketch. Hum Brain Mapp 1999; 7(3):166–94.
[17] Nichols T, Hayasaka S. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res 2003;12(5):419–46.
[18] Ardekani BA, Choi SJ, Hossein-Zadeh GA, Porjesz B, Tanabe JL, Lim KO, et al. Functional magnetic resonance imaging of brain activity in the visual oddball task. Brain Res Cogn Brain Res 2002; 14(3):347–56.
[19] Brazdil M, Dobsik M, Mikl M, Hlustik P, Daniel P, Pazourkova M, et al. Combined event-related fMRI and intracerebral ERP study of an auditory oddball task. Neuroimage 2005;26(1):285–93.
[20] Brazdil M, Mikl M, Marecek R, Krupa P, Rektor I. Effective connectivity in target stimulus processing: a dynamic causal modeling study of visual oddball task. Neuroimage 2007;35(2):827–35.
[21] Friston KJ, Holmes AP, Price CJ, Buchel C, Worsley KJ. Multisubject fMRI studies and conjunction analyses. Neuroimage 1999;10(4): 385–96.
[22] Kiebel S, Friston KJ. Anatomically informed basis functions in multisubject studies. Hum Brain Mapp 2002;16(1):36–46.
[23] Mikl M, Chlebus P, Brazdil M, Drastich A, Krupa P. Optimization of fMRI group analysis using various spatial smoothing parameters. Neuroimage 2005;26(Suppl 1, 11th Annual Meeting of The Organization of Human Brain Mapping):593.
[24] Kiebel SJ, Poline JB, Friston KJ, Holmes AP, Worsley KJ. Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. Neuroimage 1999;10(6): 756–66.
[25] Zarahn E, Aguirre GK, D'Esposito M. Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. Neuroimage 1997;5(3):179–97.
[26] Kiebel SJ, Goebel R, Friston KJ. Anatomically informed basis functions. Neuroimage 2000;11(6 Pt 1):656–67.
[27] Reimold M, Slifstein M, Heinz A, Mueller-Schauenburg W, Bares R. Effect of spatial smoothing on t-maps: arguments for going back from t-maps to masked contrast images. J Cereb Blood Flow Metab 2006; 26(6):751–9.