Weierstraß-Institut

für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

Adaptive smoothing as inference strategy: More specificity for unequally sized or neighboring regions

Marijke Welvaert[1], Karsten Tabelow[2], Ruth Seurinck[1], Yves Rosseel[1]

submitted: June 24, 2013

[1] Department of Data Analysis
Ghent University
H.Dunantlaan 1
9000 Gent, Belgium
E-Mail: marijke.welvaert@Ugent.de

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin, Germany
E-Mail: karsten.tabelow@wias-berlin.de

No. 1805

Berlin 2013

2010 *Mathematics Subject Classification.* 62P10, 62G10, 62G05 .

Key words and phrases. false positive rate and fmri and gaussian smoothing and power and structural adaptive segmentation.

This work was partially supported by the DFG Research Center MATHEON “Mathematics for key technologies” in Berlin.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

ABSTRACT. Although spatial smoothing of fMRI data can serve multiple purposes, increasing the sensitivity of activation detection is probably its greatest benefit. However, this increased detection power comes with a loss of specificity when non-adaptive smoothing (i.e. the standard in most software packages) is used. Simulation studies and analysis of experimental data was performed using the R packages **neuRosim** and **fmri**. In these studies, we systematically investigated the effect of spatial smoothing on the power and number of false positives in two particular cases that are often encountered in fMRI research: (1) Single condition activation detection for regions that differ in size, and (2) multiple condition activation detection for neighbouring regions. Our results demonstrate that adaptive smoothing is superior in both cases because less false positives are introduced by the spatial smoothing process compared to standard Gaussian smoothing or FDR inference of unsmoothed data.

1. INTRODUCTION

For many neuroscientists, spatial smoothing of fMRI data has become an automatic preprocessing step. The purpose of this smoothing procedure can be threefold. First, spatial smoothing moderates intersubject variation in brain anatomy, especially when individual brains are transformed to a standard brain space in order to allow intersubject comparison. Second, the smoothing of fMRI data will increase the signal-to-noise ratio (SNR). Third, a voxel-based mass-univariate analysis of fMRI data calls for the need of multiple testing corrections. In this context, spatial smoothing enhances random field theory (RFT) based inference (see for example Worsley, 2003).

Although spatial smoothing, also referred to as spatial filtering, is mostly performed during the preprocessing stage of the analysis, it is actually a crucial step in the whole data analysis process because of its influence on the sensitivity and specificity of the activation detection analysis. The critical point here is that we need a useful estimate of the required width of smoothing. Depending on the goal of smoothing, guidelines vary substantially. To allow intersubject averaging, more smoothing might be necessary (e.g. a Gaussian kernel full-width half-maximum (FWHM) of 8 mm, which is often the default value in software packages) (Friston et al, 2007). In the context of increasing SNR and based on the matched filter theorem (Rosenfeld and Kak, 1982), one wants to take into account the size of the region of interest (ROI), which can vary roughly from 2 to 10 mm, or even more. Finally, the spatial smoothness needed for valid statistical inference can be rather small (around 4 mm will mostly be sufficient) (Friston et al, 1994). In terms of maintaining spatial structure, it would be better to smooth as little as possible. A smoothing kernel width that is at least twice the voxel size should be appropriate in almost all cases (Poldrack et al, 2011).

1.1. **Non-adaptive smoothing.**

Spatial smoothing is applied by the convolution of each volume of the fMRI dataset and a Gaussian kernel. In practice this translates to the signal in the volume being blurred by averaging the data over all voxels in a spherical region. Based on the size of the FWHM of the Gaussian kernel, local weights determine if a certain voxel is part of this region. In the case of non-adaptive smoothing, these weights are acting as indicators on how far the smoothing kernel reaches in space and are completely defined by the value of the FWHM.

Although spatial smoothing increases the sensitivity of activation detection, it also has a major disadvantage, namely the decrease of spatial resolution. This results in loss of information on the spatial extent and the shape of activated regions. Moreover, when the smoothing kernel is large compared to the activated area, the sensitivity of activation detection will decrease and false positives are introduced. Some algorithms were developed to overcome this increase of false positive rate and decrease of spatial resolution, such as, for example, scale space methods (Poline and Mazoyer, 1994) and non-linear filtering (Descombes et al, 1998). In this paper, we will focus on the most standard spatial smoothing method, referred to as Gaussian smoothing, as an example of non-adaptive smoothing.

1.2. **No smoothing.** In a reaction to the specificity loss caused by non-adap-tive smoothing, some researchers omit the smoothing step and perform inference based on unsmoothed data. The result is that activation detection is probably highly specific, but on the other hand might also be overly conservative. First, there has to be enough signal/contrast in the data to be sensitive to the activation, and second, inference that corrects for multiple comparisons based on RFT will be conservative due to unsufficient smoothness in the data (Logan and Rowe, 2004). One solution to get more sensitive results out of unsmoothed data would be to use False Discovery Rate control (FDR) (Benjamini and Hochberg, 1995), since this method is known to be more sensitive (Logan and Rowe, 2004). We will consider both strategies as alternatives to spatial smoothing of the data.

1.3. **Adaptive smoothing.** A second option to overcome the drawbacks of non-adaptive spatial smoothing, and probably also of no smoothing, is to smooth adaptively. With this smoothing technique, the smoothing width is not chosen a priori but based on the data by using for example adaptive region growing, nonstationary spatial Gaussian Markov random fields or adaptive weights smoothing (e.g. Lu et al, 2003; Harrison et al, 2008; Yue et al, 2010; Tabelow et al, 2006; Polzehl et al, 2010). We will discuss two of these methods in more detail which will serve as examples of adaptive smoothing.

In 2006, Tabelow et al. introduced a structural adaptive smoothing procedure based on the propagation-separation approach (Polzehl and Spokoiny, 2006). Without using any prior anatomical knowledge, the methodology takes into account the size and shape of activated regions (Tabelow et al, 2006). During an iterative process adaptive weighting schemes at each location are determined based on the parameter estimates of the statistical parametric map (SPM). These weights separate areas of distinct parameter values in particular activated voxels from non-activated voxels. This avoids the blurring bias at the borders of these areas typically observed in non-adaptive Gaussian filtering (see Tabelow et al, 2006, for the technical details). Signal detection can then be performed, for instance, using thresholds based on Random Field Theory (Adler, 2000; Worsley, 1994, 2003). Additionally, Polzehl et al (2010) presented a new version of the structural adaptive smoothing algorithm, namely structural adaptive segmentation. Their extended algorithm combines the estimation and smoothing step with the inference step based on multiscale tests. Both simulations and analysis of real data showed higher sensitivity and specificity compared to the original adaptive smoothing procedure (see Polzehl et al, 2010, for the technical details).

A major advantage of these adaptive smoothing techniques is that users only have to provide the maximum bandwidth of smoothing while the algorithm determines the optimal local weighting scheme. Moreover, when no spatial structure can be detected based on the functional activation, the smoothing procedure reduces to non-adaptive Gaussian smoothing such that the SPM under the null hypothesis is again approximately a Gaussian Random Field.

It should be noted that spatial smoothing is usually part of the preprocessing pipeline. However, parameter estimation, with the exception of prewhitening effects, is not effected by the spatial smoothing and therefore the order of smoothing and parameter estimation is interchangeable (Tabelow et al, 2006). Since the adaptive smoothing procedures rely heavily on the SNR in the data, parameter estimation before smoothing is necessary as a variance and dimension reduction step. Therefore, in this paper, all smoothing procedures will be applied after parameter estimation.

1.4. **Current study.** Despite the fact that the disadvantages of non-adaptive spatial smoothing are well-known and that new improved methods (e.g. adaptive smoothing) have been introduced, Gaussian smoothing is still used almost exclusively, probably because it is widely available. In this paper, we address two particular situations that are applicable to fMRI research and for these situations we provide a systematic analysis of the dependence of the sensitivity and specificity on smoothing parameters in a combination with inference methods. By applying spatial smoothing, we will increase the SNR resulting in higher sensitivity, but, by comparing the results for different smoothing and inference methods, we will look for an optimal trade-off between gaining sensitivity and losing specificity.

In the first study (section 2), experimental tasks are considered that cause activation in multiple unequally sized regions. When we want to take into account the spatial extent of these regions, the choice of the value of the smoothing kernel width is non-trivial. Assuming that the actual size of the activated regions is known (e.g. based on anatomical structures, localizer tasks or previous analyses of similar tasks), the researcher is confronted with a choice of the FWHM value ranging from the size of the smallest region to the size of the largest region. The former will result in undersmoothing the larger regions, while choosing the latter will oversmooth the smaller regions. The second study (section 3) investigates the effect of spatial smoothing on the specificity and sensitivity of activation detection when two neighbouring regions are activated due to different tasks, conditions or contrasts. Here, we definitely want to avoid oversmoothing because this could result in an unintended overlap between the activation regions.

For both studies, results for unsmoothed, non-adap-tively smoothed and adaptively smoothed data are contrasted. For the unsmoothed data, RFT based Family-Wise Error (FWE) control and FDR control[1] are compared for two reasons. First, since the latter inference method is known to be more sensitive (Logan and Rowe, 2004), we will investigate its properties as an alternative to using spatial smoothing to augment the power. Secondly, due to unsufficient smoothness, RFT inference is known to be too conservative on unsmoothed data. For the smoothed data, inference is based on RFT (except for the adaptive segmentation algorithm which used multiscale tests). Other inference methods, like for example cluster-based inference are not used here as they have the same blurring problem as RFT when relying on a Gaussian filter. Using more realistic simulations (as opposed to Tabelow et al, 2006; Polzehl et al, 2010) the empirical power and false positive rate (FPR) are assessed at several contrast-to-noise ratio (CNR) values (section 2.1 and 3.1). As an illustration, the conditions from both studies are also applied on real data (section 2.2 and 3.2).

2. STUDY 1: REGIONS WITH UNEQUAL SIZE

2.1. **Simulation study.**

2.1.1. *Data generation and design.* We simulated fMRI datasets of size $20 \times 20 \times 20 \times 107$ consisting of 2 mm^3 isotropic voxels in R (R Development Core Team, 2010) using the package **neuRosim** (Welvaert et al, 2011). The stimulus function was based on a block design with 3 activation blocks that each lasted 15 scans (see Figure 1 for a display of the timecourse). We modelled 2 activated regions (spheres with diameter 6mm and 10mm respectively, see Figure 1). *Rich* noise was added including temporal correlations ($\rho = 0.3$), spatial ($\rho = 0.7$) correlations and physiological noise (i.e. noise due to heart rate, respiratory rate and task-related noise). Specifically, the noise consisted for 10% of white noise, 30% temporally correlated noise, 20% low frequency drift, 10% physiological noise, 10% task-related noise and 20% spatially correlated noise (see Welvaert and Rosseel, 2012, for more details on the noise generation). These raw data were then analysed in R using the package **fmri** (Tabelow and Polzehl, 2011). Six smoothing conditions were considered: (**Condition 1**) no spatial smoothing, (**Condition 2a**) Gaussian smoothing with the size of the smoothing kernel equal to the size of the smallest region (FWHM=6 mm), (**Condition 2b**) Gaussian smoothing with the size of the smoothing kernel equal to the average size of the regions, corresponding to the default value (FWHM=8 mm), (**Condition 2c**) Gaussian smoothing with the size of the smoothing kernel equal to the size of the largest regions (FWHM=10 mm), (**Condition 3**) structural adaptive smoothing with maximum bandwidth FWHM=10 mm, and (**Con-dition 4**) structural adaptive segmentation with maximum bandwidth FWHM=10 mm. We also varied the CNR level (i.e., ratio of changes in the signal due to the experiment and fluctuations due to noise) of the data between 0.02 and 0.5, and each dataset was replicated 100 times. For each replication, we assessed the power and

[1] We explicitly discard cluster-based FDR methods (e.g. Chumbley and Friston, 2009) because they are also dependent on RFT inference.

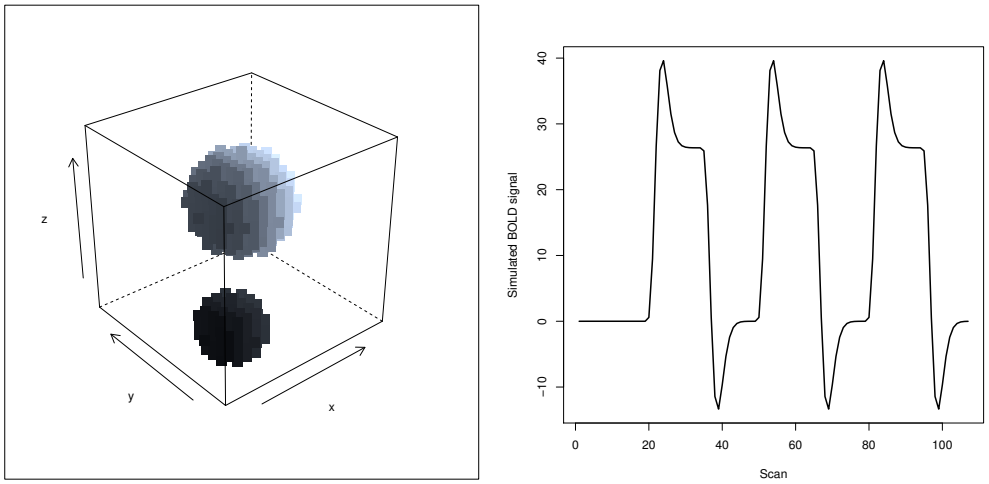

FIGURE 1. Ground truth activation in simulation study 1. Left: location of the activated regions (grayscale values reflect position in space). Right: timecourse of the block design.

TABLE 1. Overview of the spatial smoothing and multiple corrections methods applied in simulation study 1.

Condition	Smoothing	Inference
1a	none	Random Field Theory
1b	none	False Discovery Rate
2a	non-adaptive (FWHM = 6 mm)	Random Field Theory
2b	non-adaptive (FWHM = 8 mm)	Random Field Theory
2c	non-adaptive (FWHM = 10 mm)	Random Field Theory
3	adaptive smoothing	Random Field Theory
4	adaptive segmentation	Multiscale Tests

FPR of the detected activation based on a general linear model analysis including an AR(1) temporal correlation model conducted using **fmri** (Tabelow and Polzehl, 2011).

The resulting statistical parametric maps (SPMs) were smoothed according to the six smoothing conditions. So, in this context spatial smoothing is not considered to be part of the pre-processing steps (Tabelow et al, 2006). Sensitivity was measured by means of the average power, which was calculated as the ratio of correctly detected voxels and the total number of active voxels. Similarly, specificity was measured as the average false positive rate (FPR) obtained by taking the ratio of falsely detected voxels and the total number of non-active voxels. All results are corrected for multiple comparisons at $p < 0.05$ based on a Family-wise Error (FWE) correction using Gaussian Random Field theory in the no smoothing, non-adaptive smoothing, and adaptive smoothing conditions (Worsley, 2003), and using multiscale tests for the adaptive segmentation condition (Dümbgen and Spokoiny, 2001; Polzehl et al, 2010). As a comparison and because RFT might not work properly on unsmoothed data due to unsufficient smoothness, we also applied FDR based inference (Benjamini and Hochberg, 1995) on the non-smoothed data. Table 1 provides an overview of the combination of the smoothing conditions and the multiple corrections methods.

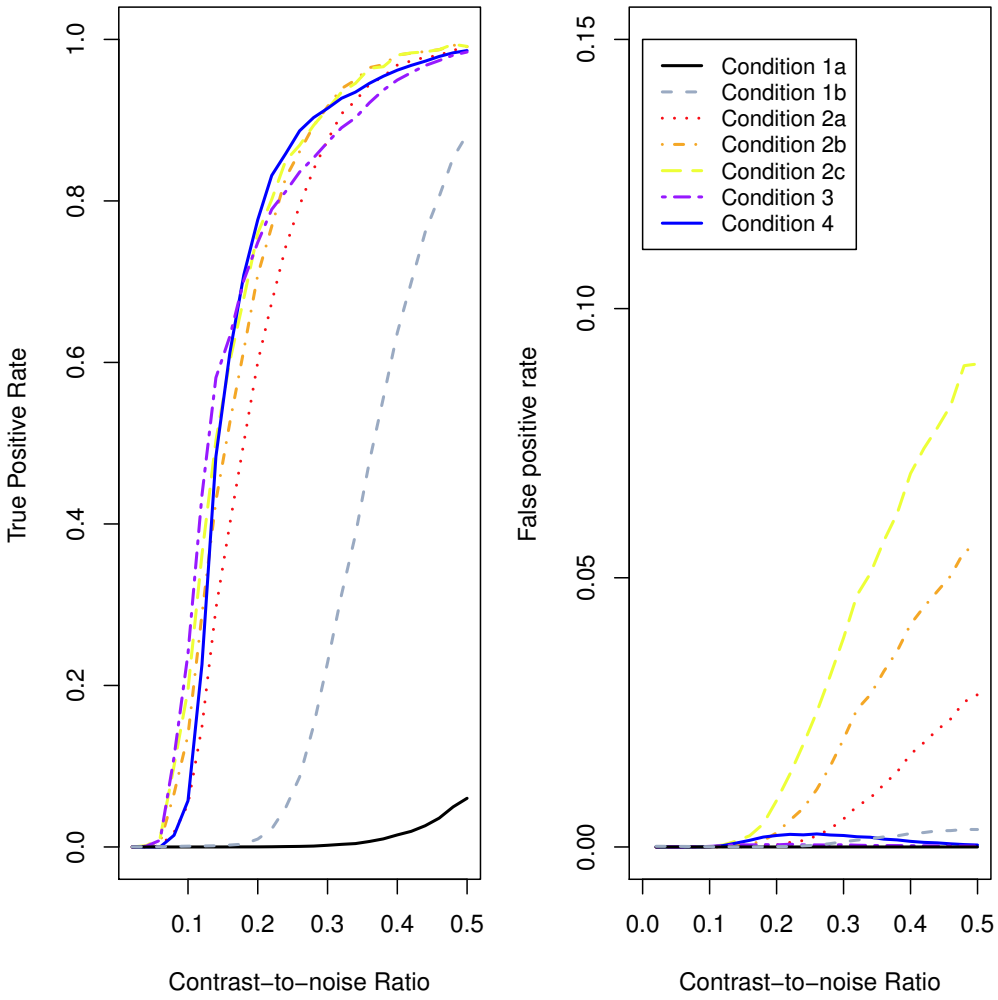

FIGURE 2. Power (left) and FPR (right) results for the conditions in simulation study 1. Table 1 presents an overview of the conditions.

2.1.2. *Results.* The results are presented in Figure 2. The power results on the left hand side of the figure and the FPR results on the right hand show that there is indeed a difference between the smoothing conditions. In the no smoothing condition combined with RFT inference, the overall power and FPR levels are very low. This might be due to unsufficient smoothness in the data for the algorithm to work properly. By comparison, using FDR control on the unsmoothed data already gives more power, so this will be our baseline condition to compare the smoothing methods against. By spatially smoothing the data we would like to obtain higher power, however, ideally, the FPR should stay as low as when no smoothing is applied. Looking at the power results first (Figure 2, left panel), we observe higher power values for all smoothing methods as expected. The differences between the conditions are rather small, although it seems that adaptive segmentation results in the highest sensitivity levels, while Gaussian smoothing with the smallest kernel (FWHM = 6 mm) provide a lower bound on the power results.

So, while the obtained power results are quite similar, more striking differences are observed for the FPR results (figure 2, right panel). For the Gaussian smoothing conditions, more false positives are observed with increasing kernel width and increasing CNR values. Sin-ce there are almost no false positives when no smoothing is applied, the increased FPR rate observed in the Gaussian smoothing conditions is completely attributable to the smoothing procedure. For example, with a medium-to-high CNR of 0.4,

about 5% false positives are introduced by the Gaussian smoothing procedure in the case the default value of the FWHM is used (i.e. 8 mm). In contrast, the FPR results for the adaptive smoothing and adaptive segmentation techniques are very similar to the no smoothing condition; the number of false positives is very low. Also in the case of no smoothing with FDR based inference, the number of false positives appears to be well controlled.

Based on these simulation results, we can conclude that when there are multiple activated regions, higher sensitivity can be obtained by either adaptive smoothing, adaptive segmentation or Gaussian smoothing, compared to no smoothing (combined with either RFT or FDR). However, only the adaptive procedures (adaptive smoothing and adaptive segmentation) succeed in maintaining the specificity while this greatly decreases for the Gaussian smoothing conditions.

2.2. **Real data.** We applied the smoothing procedures as described in the simulation study above to experimental fMRI data from a passive viewing task used to localize hV5/MT+ involved in motion perception (Seurinck et al, 2011). The data were acquired on a 3T Siemens TRIO MR scanner (Siemens Medical Systems, Erlangen, Germany) with scanning parameters TR $= 1940$ ms, TE $= 35$ ms, flip angle $= 80^{\circ}$, 28 slices, slice thickness $= 3$ mm with a distance factor of 17%, FOV $= 244$ mm, matrix $= 64 \times 64$. We used 425 EPI images (corresponding to the first run) from a random subject in the dataset.

All images were motion corrected[2] and normalized to the MNI152 template using SPM8. Spatial smoothing was performed using the R package **fmri** (Tabelow and Polzehl, 2011) corresponding to the same conditions as in simulation study 1. In a GLM analysis, we tested the contrast *moving stimuli* versus *static stimuli*. We expect larger activation in hV5/MT+ for the moving stimuli compared to the static stimuli. Figure 3 shows the results for the different smoothing conditions. Since locator tasks have typically high CNR, the results for the no smoothing condition (Figure 3a) already show clear bilateral activation in hV5/MT+. Based on the activation detected in this condition, we distinguish two activated regions, one in the right hemisphere, which we will refer to as the small region, and one in the left hemisphere, which we will refer to as the large region. The size of the FWHM values for the spatial smoothing are in this case based on the sizes of the detected regions in these unsmoothed data. When using FDR controlled inference on the unsmoothed data (Figure 3b), we observe more sensitive results, but surpisingly also more activation outside the hV5/MT+ region which is improbable given the contrast. Non-adaptive (Gaussian) smoothing with a kernel width equal to the small region (FWHM = 6) gives slightly extended activation regions (Figure 3c). However, increasing the kernel width further results in loss of detected activation. For Gaussian smoothing with FWHM = 8 mm, the activation in the right hemisphere is greatly diminished (Figure 3d), and with FWHM = 10 mm, it disappears completely (Figure 3e). For this kernel width, also in the left hemisphere only half of the activation survives the threshold. In contrast, the adaptive smoothing techniques (Figure 3f-g) can produce slightly larger hV5/MT+ activations with increased sensitivity compared to the no smoothing condition. This demonstrates that loss of activation due to oversmoothing is overcome with adaptive smoothing. Additionally, while sensitivity to activation is increased for both adaptive procedures, the highest sensitivity levels are observed for adaptive segmentation. Concerning the extra activation outside the hV5/MT+ regions that was detected in the unsmoothed - FDR condition, spatial smoothing (all methods) succeeds in decreasing this activation, resulting in clearly localized activation areas.

3. Study 2: neighbouring regions

3.1. **Simulation study.**

[2]The influence of spatial smoothing in this stage was minimized by setting the FWHM of the Gaussian smoothing kernel to 1 mm, before estimating the realignment parameters.

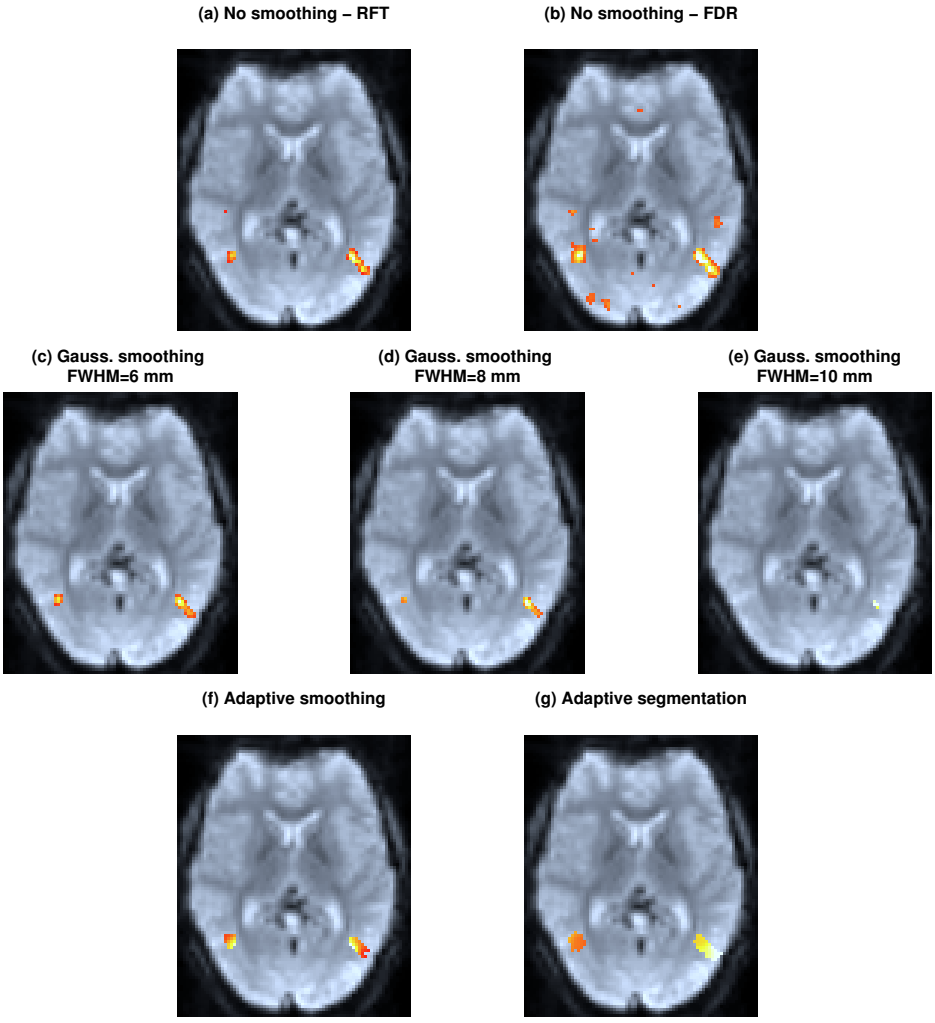

FIGURE 3. Example slices (Left=Right orientation) showing hV5/MT+ activation in a visual localizer experiment (Seurinck et al, 2011). Brighter colours indicate stronger activation (based on the estimated signal of active voxels in case of adaptive segmentation, and on the p-values otherwise).

3.1.1. *Data generation and design.* We simulated fMRI datasets of size $18 \times 18 \times 18 \times 120$ consisting of 2 mm^3 isotropic voxels in R (R Development Core Team, 2010) using the package **neuRosim** (Welvaert et al, 2011).

The experiment was a block design with 2 conditions. Both conditions contain 5 activation blocks of 10 scans which are alternately active. We modelled 2 activated regions (6×6 cubes) next to each other accounting for activation based on condition 1 and 2, respectively (see Figure 4), and we used the same noise model as in simulation study 1. We considered 4 conditions: (1) no smoothing with RFT inference, (2) no smoothing with FDR inference, (3) Gaussian smoothing with FWHM=6 mm (i.e. congruent with the size of the activated region), and (4) adaptive segmentation with the maximal bandwidth equal to 6 mm. The CNR varied between 0.02 and 1, which are values common in fMRI research, and we simulated 100 replications of each dataset. The data were analyzed using the R package **fmri** (Tabelow and Polzehl, 2011) and the power and FPR were determined for each activation condition separately.

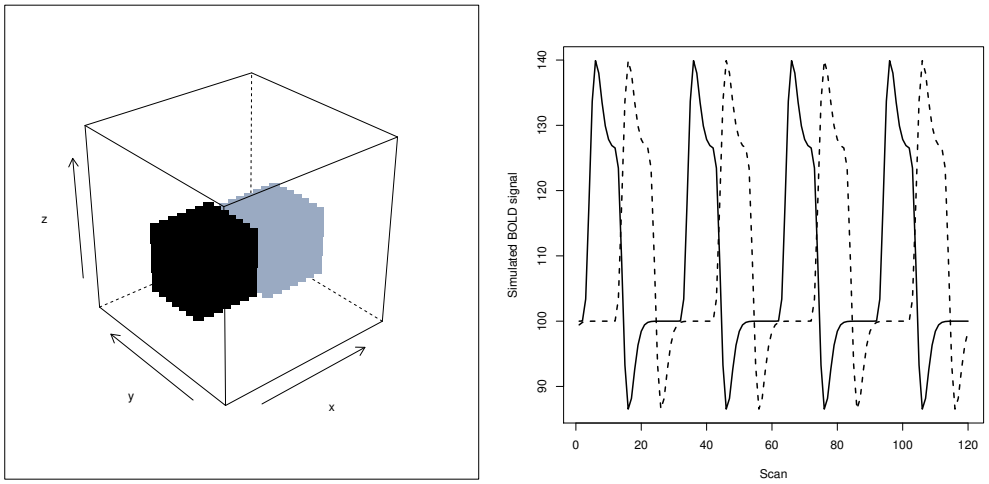

FIGURE 4. Ground truth activation in simulation study 2. Left: location of the activated regions. Right: timecourses of the block design.

3.1.2. *Results.* The results are shown in Figure 5. As in simulation study 1, in the no smoothing condition using RFT inference, both power and FPR are very low for all CNR levels. Again inference based on RFT might not be entirely correct because of unsufficient smoothness. The FDR based results show indeed more power, while keeping the FPR low. This condition will serve again as our baseline. Looking at the power results (figure 5, left panel), we see increasing power with increasing CNR levels for both smoothing methods and in both conditions. For the tests of both conditions, Gaussian smoothing results in slightly more power compared to adaptive segmentation for lower CNR values, but the reverse is the case for higher CNR values. Overall, the differences in sensitivity between both smoothing techniques are rather low. The FPR results (figure 5, right panel) demonstrate a better performance of adaptive segmentation. This smoothing procedure can control the number of false positives very well in both conditions. However, for the Gaussian smoothing condition, the FPR increases with increasing CNR values. Based on these simulation results we would again recommend adaptive smoothing as the better method to increase the sensitivity while maintaining specificity in the case of neighbouring regions.

3.2. **Real data.** Similar to the first simulation study, we demonstrate the impact of the smoothing procedures on the results from experimental fMRI data for the case of neighbouring regions. We used the same data as in the previous example (Seurinck et al, 2011), however now we focus on different contrasts. In a GLM analysis we tested the following two contrasts: (1) motion versus rest, and (2) static versus rest. These are again localizer contrasts indicated to find seperately active regions based on moving stimuli on the one hand and static stimuli on the other hand. We do not expect any overlap between the regions based on these contrasts. Similar to the second simulation study, we consider four conditions: (1) no smoothing with RFT inference, (2) no smoothing with FDR inference, (3) Gaussian smoothing with the FWHM matched to the size of the activated region (i.e. 8 mm), and (4) adaptive segmentation. Figure 6 shows results for the different smoothing procedures on the most representative slice. When the data are not smoothed (upper panel), RFT inference results in clearly distinct activation clusters for both conditions without any overlap. However, using FDR control shows a very different activity pattern. The static stimuli contrast (yellow) results in widespread activation in several areas and also a lot of overlap (green) is detected with the moving stimuli contrast (blue). Smoothing the data (lower panel) results in two clusters that show small overlap in the Gaussian smoothing condition, despite the fact that the value of the FWHM is slightly smaller than the size of the region (based on the unsmoothed data). In contrast, using adaptive

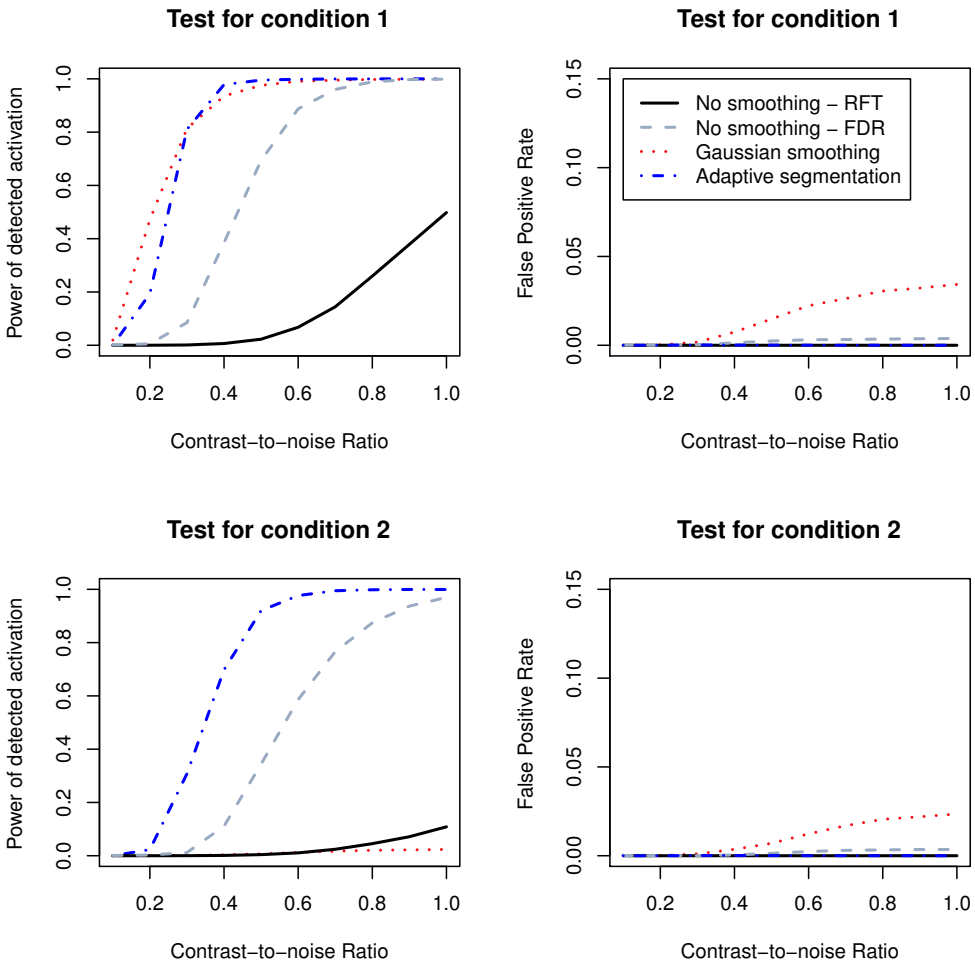

FIGURE 5. Power (left panel) and FPR (right panel) results for each condition in simulation study 2.

segmentation enables us to localize two activation regions for the moving and static stimuli separately without any overlap.

Next to more specific activation detection in the adaptive segmentation condition, the sensitivity for activation strength is also higher in this condition. Figure 7 shows the probability density functions of the p values of active voxels for the non-smoothing and Gaussian smoothing conditions and the scaled estimated effect for the adaptive segmentation condition. Based on this figure, we can conclude that for both estimated contrasts the activation evidence is highest for the adaptive segmentation, followed by the unsmoothed data with FDR control.

As already demonstrated in the simulations, this real data example again indicates that adaptive segmentation is successful in maintaining the original shape of the activated region (i.e., increasing specificity compared to Gaussian smoothing), while increasing the overall sensitivity of activation detection at the same time.

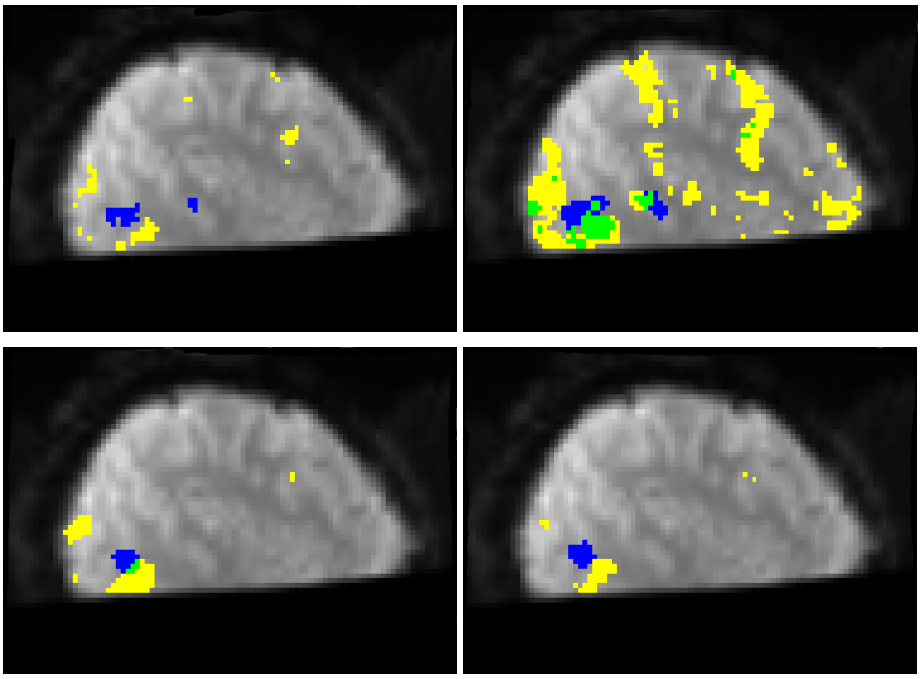

FIGURE 6. Application of the smoothing conditions from simulation study 2 to the visual localizer data from Seurinck et al (2011). Upper left: no smoothing (RFT); upper right: no smoothing (FDR); lower left: Gaussian smoothing; lower right: adaptive segmentation. Blue indicates activation for moving stimuli, yellow codes for activation for static stimuli and green shows overlap between the two contrasts.

4. DISCUSSION

Despite the consensus on the benefits of spatial smoothing, the procedure is often used in a standard preprocessing pipeline using default values, especially with regard to the kernel width. For example, only 8% of studies that report using spatial smoothing, describe the reason why a particular value of smoothing kernel width is chosen (Carp, 2012). In case a justification is provided, the relation with the specific selected smoothing kernel is at least vague. Poldrack et al (2011) give as a guideline that a smoothing kernel width of at least twice the voxel size would be appropriate in many cases (pp. 51–52). A second problem is that applying non-adaptive spatial smoothing comes at a cost of losing specificity. To accommodate this issue, the kernel width should be chosen wisely to achieve the desired increase in sensitivity while maintaining an acceptable specificity level. Given these confusing guidelines, the drawbacks of standard smoothing techniques and constant development of more advanced analyses, a systematic evaluation of the impact of spatial smoothing is called for.

In this paper we focused on two specific scenarios often encountered in fMRI research, namely (1) activation in unequally sized regions and (2) activation in neighbouring regions. For both cases several methods for smoothing and signal detection were systematically evaluated using realistic simulations. Based on the power results, we can conclude that all spatial smoothing procedures are successful in increasing the power of the activation detection analysis. Logan and Rowe (2004) have already demonstrated on unsmoothed that FDR based inference is more sensitive compared to RFT inference, however the sensitivity levels of the FDR method are not as high when compared to the smoothed data results. Although increased sensitivity is desirable, our FPR analyses have shown that differences between smoothing methods are quite dramatic for the specificity results. Applying non-adaptive smoothing will inevitably lead to more false positives. In the case of unequally sized regions, the chosen value of the FWHM can

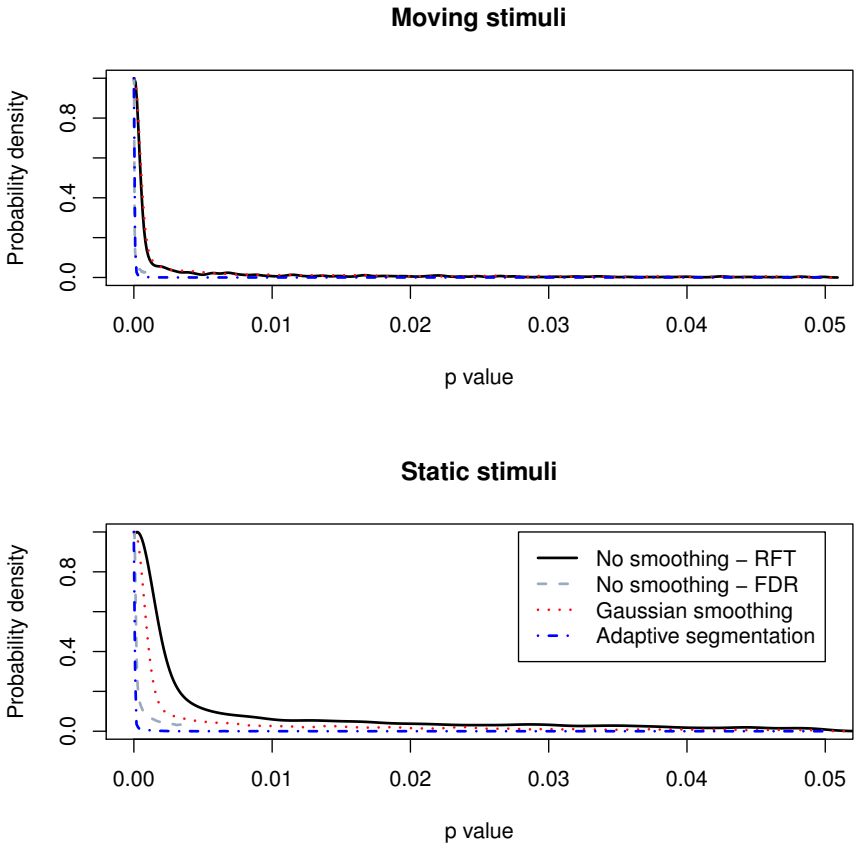

FIGURE 7. Distribution of the p values of active voxels for moving stimuli and static stimuli respectively. Note that for the adaptive segmentation condition the scaled estimated effect is plotted instead of a p value.

even increase this effect further. On the other hand, even when the FWHM can be matched to the size of the activated region, an overlap will be created between neighbouring regions. Both effects were not present when applying adaptive smoothing. Almost no false positives are introduced by these smoothing methods.

A demonstration on real experimental data further accords with these results. However, it should be noted that, while the FDR based inference of unsmoothed data in the simulation studies could control the FPR, the same analysis of the real data showed more activation clusters that seemed improbable given the localizer contrasts. Applying FDR inference on smoothed data might be a possible solution to control this unwanted activation, but the method requires independent tests to be valid which is not the case for smoothed data (although Benjamini and Yekutieli (2001) developed a correction for some forms of dependency) or has to rely oRFT to make cluster-based inference (Chumbley and Friston, 2009).

With respect to the inference methods, we saw that, confirming the results of Polzehl et al (2010), adaptive smoothing combined with multiscale tests (i.e. adaptive segmentation) results in higher sensitivity levels compared to using RFT based inference. We focused on voxelwise inference using RFT corrected thresholds but we expect that the results for other inference methods, such as cluster-based techniques, would be similar because they would suffer from the same blurring effect induced by non-adaptive smoothing.

Similarly, other adaptive smoothing techniques that are not considered here (e.g. Lu et al, 2003; Harrison et al, 2008; Yue et al, 2010), but also make use of the data to adaptively smooth the images, are expected to behave in the same way. By using knowledge on shape and extent of the activated region that is present in the data, these more advanced techniques will provide more reliable results with respect to the obtained specificity.

By specifically focusing on two cases applicable to single-subject fMRI, we left the question of the impact of spatial smoothing on higher-level analyses open for the time being. We used spatial smoothing in the first place as a method to increase the SNR of the data, but in higher-level analyses spatial smoothing is also used to make intersubject comparison more feasible. It is possible that some degree of specificity loss is necessary to enable enough overlap between different subjects but this will be a subject for future research.

5. Conclusions

More advanced adaptive smoothing procedures can be used as an inference strategy to obtain more specificity for example with unequally sized or neighbouring regions. Compared to the widely used non-adaptive spatial smoothing, adaptive smoothing has the advantage of controlling the number of false positives while increasing the power. In addition, using adaptive segmentation methods always includes a built-in multiple comparisons correction based on multiscale tests.

6. Information Sharing Statement

The R packages **neuRosim** and **fmri** are open-source packages and can be downloaded and installed from CRAN (http://cran.r-project.org).
SPM8 is a Matlab toolbox that is freely available from http://www.fil.ion.ucl.ac.uk/spm/software/spm8/.
The data used in the real data example are available upon request.

7. Conflict of interest

The authors declare that they have no conflict of interest.

References

Adler R (2000) On excursion sets, tube formulae, and maxima of random fields (special invited paper). Annals of Applied Probability 10:1–74

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Statist Soc Ser B 57:289–300

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. The Annals of Statistics 29:1165–1188

Carp J (2012) The secret lives of experiments: Methods reporting in the fMRI literature. NeuroImage 63:289–300

Chumbley J, Friston K (2009) False discovery rate revisited: Fdr and topological inference using gaussian random fields. NeuroImage 44:62–70

Descombes X, Kruggel F, von Cramon D (1998) Spatio-temporal fMRI analysis using markov random fields. IEEE Transactions on Medical Imaging 17:1028–29

Dümbgen L, Spokoiny V (2001) Multiscale testing of qualitative hypotheses. Annals of Statistics 29:124–152

Friston K, Worsley K, Frackowiak R, Mazziotta J, Evans A (1994) Assessing the significance of focal activations using their spatial extent. Human Brain Mapping 1:214–220

Friston K, Ashburner J, Kiebel S, Nichols T, Penny W (eds) (2007) Statistical Parametric Mapping: The Analysis of Functional Brain Images. Academic Press, Massachussets, USA

Harrison L, Penny W, Daunizeau J, Friston K (2008) Diffusion-based spatial priors for functional magnetic resonance images. NeuroImage 41:408–423

Logan B, Rowe D (2004) An evaluation of thresholding techniques in fMRI analysis. NeuroImage 22:95–108

Lu Y, Jiang T, Zang Y (2003) Region growing method for the analysis of functional mri data. NeuroImage 20:455–465

Poldrack R, Mumford J, Nichols T (2011) Handbook of functional MRI data analysis. Cambridge University Press, New York, USA

Poline J, Mazoyer B (1994) Enhanced detection in brain activation maps using a multifiltering approach. Journal of Cerebral Blood Flow Metabolism 14:639–642

Polzehl J, Spokoiny V (2006) Propagation-separation approach for local likelihood estimation. Probability Theory and Relative Fields 135:335–362

Polzehl J, Voss H, Tabelow K (2010) Structural adaptive segmentation for statistical parametric mapping. NeuroImage 52:515–523

R Development Core Team (2010) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL `http://www.R-project.org`, ISBN 3-900051-07-0

Rosenfeld A, Kak A (1982) Digital picture processing 2. Academic Press, Orlando, Florida

Seurinck R, de Lange F, Achten E, Vingerhoets G (2011) Mental rotation meets the motion aftereffect: The role of hV5/MT+ in visual mental imagery. Journal of Cognitive Neuroscience 23:1395–1404

Tabelow K, Polzehl J (2011) Statistical parametric maps for functional MRI experiments in r: The package fmri. Journal of Statistical Software 44:1–21

Tabelow K, Polzehl J, Voss H, Spokoiny V (2006) Analyzing fMRI experiments with structural adaptive smoothing procedures. NeuroImage 33:55–62

Welvaert M, Rosseel Y (2012) How ignoring physiological noise can bias the conclusions from fMRI simulation studies. Journal of Neuroscience Methods 211:125–132

Welvaert M, Durnez J, Moerkerke B, Verdoolaege G, Rosseel Y (2011) neuRosim: An R package for generating fMRI data. Journal of Statistical Software 44:1–18

Worsley K (1994) Local maxima and the expected euler characteristic of excursion sets of χ^2, f and t fields. Advances in Applied Probability 26:13–42

Worsley K (2003) Detecting activation in fmri data. Statistical Methods in Medical Research 12:401–418

Yue Y, Loh J, Lindquist M (2010) Adaptive spatial smoothing of fmri images. Statistics and Its Interface 3:3–13