Journal of Cerebral Blood Flow & Metabolism (2005), 1–9

www.jcbfm.com

Effect of spatial smoothing on *t*-maps: arguments for going back from *t*-maps to masked contrast images

Matthias Reimold[1], Mark Slifstein[2], Andreas Heinz[3], Wolfgang Mueller-Schauenburg[1] and Roland Bares[1]

[1]*Department of Nuclear Medicine, University of Tuebingen, Tuebingen, Germany;* [2]*Division of Functional Brain Mapping, Columbia University and New York State Psychiatric Institute, New York, New York, USA;* [3]*Department of Psychiatry and Psychotherapy, Charité (CCM), Humboldt University of Berlin, Berlin, Germany*

Voxelwise statistical analysis has become popular in explorative functional brain mapping with fMRI or PET. Usually, results are presented as voxelwise levels of significance (*t*-maps), and for clusters that survive correction for multiple testing the coordinates of the maximum *t*-value are reported. Before calculating a voxelwise statistical test, spatial smoothing is required to achieve a reasonable statistical power. Little attention is being given to the fact that smoothing has a nonlinear effect on the voxel variances and thus the local characteristics of a *t*-map, which becomes most evident after smoothing over different types of tissue. We investigated the related artifacts, for example, white matter peaks whose position depend on the relative variance (variance over contrast) of the surrounding regions, and suggest improving spatial precision with 'masked contrast images': color-codes are attributed to the voxelwise contrast, and significant clusters (e.g., detected with statistical parametric mapping, SPM) are enlarged by including contiguous pixels with a contrast above the mean contrast in the original cluster, provided they satisfy $P < 0.05$. The potential benefit is demonstrated with simulations and data from a [^{11}C]Carfentanil PET study. We conclude that spatial smoothing may lead to critical, sometimes-counterintuitive artifacts in *t*-maps, especially in subcortical brain regions. If significant clusters are detected, for example, with SPM, the suggested method is one way to improve spatial precision and may give the investigator a more direct sense of the underlying data. Its simplicity and the fact that no further assumptions are needed make it a useful complement for standard methods of statistical mapping.
Journal of Cerebral Blood Flow & Metabolism advance online publication, 5 October 2005; doi:10.1038/sj.jcbfm.9600231

Keywords: fMRI; PET; smoothing artifacts; statistical parametric mapping

Introduction

Voxelwise statistical analysis has become popular in explorative functional brain mapping (Ashburner *et al*, 2003) and powerful tools for spatial normalization (Ashburner and Friston, 1999) and statistical analysis (Friston *et al*, 1995), including correction for multiple testing (Worsley *et al*, 1992; Friston *et al*, 1994) are publicly available. In the absence of an anatomically defined *a priori* hypothesis, statistical tests can be calculated for each voxel after spatial normalization and smoothing. Spatial smoothing is required to cope with interindividual functional anatomic variability that is not compensated by spatial normalization, and to improve the signal-to-noise ratio. Following the matched filter theorem (Rosenfeld and Kak, 1982) that states that the optimal smoothing kernel should match the signal to be detected, smoothing kernels of up to a *FWHM* (full-width at half-maximum) of 20 mm have been used. Smoothing with an *FWHM* of 10 to 15 mm is common in PET studies; often smaller kernels are applied in fMRI. The resulting *t*-maps (t = estimated parameter divided by its standard error) are masked at a certain threshold of voxel-level significance (often $P < 0.001$), and depicted as an overlay on corresponding anatomic sections, as maximum intensity projections or as surface projections. Those clusters that survive correction for multiple testing are frequently characterized by the

Correspondence: Dr M Reimold, PET, Roentgenweg 15, 72076 Tuebingen, Germany.
E-mail: matthias.reimold@uni-tuebingen.de
This work was supported in part by the Deutsche Forschungsgemeinschaft (Ba 1026/6-2 and He 2597/7-2).
Received 17 March 2005; revised 4 July 2005; accepted 22 August 2005

coordinates of the maximum *t*-value and the associated 'nearest gray matter'. Many of these techniques have originally been developed to detect cortical patterns of neural activation (e.g., with $[^{15}O]H_2O$ PET or fMRI BOLD signal) and are now increasingly applied also in neuroreceptor/-transporter studies.

It is widely accepted that smoothing limits spatial selectivity and significant clusters are to be interpreted as *regional* results rather than anatomically precise information. However, little attention is being given to the fact that smoothing affects *t*-maps differently from 'contrast images' (contrast = linear combination of parameter estimates in the general linear model that is supposed to reflect the interesting physiologic parameter). While the latter change with *FWHM* in a way one might guess intuitively, *t*-statistics formed from the smoothed data are also affected by the nonlinear interaction of the filter kernel with the voxel variances, which becomes most evident after smoothing over different types of tissue (e.g., gray matter, white matter) with a different variability. Voxelwise maps of parameters such as binding potential have two main sources of variance—measurement error (all sources of error associated with instrumentation, operator imprecision and the stochastic nature of isotope decay) and authentic between-subject physiologic variance. The physiologic variance can also be decomposed into two main sources—variance from nuisance variables such as intersubject differences in nonspecific binding, and differences in receptor availability, the variable of interest. The variance attributable to between-subject differences in receptor availability can be considerable, leading to larger variance in gray matter tissues with high receptor density compared with other tissues, especially white matter. If, following smoothing, voxel by voxel statistical tests are performed, spatially inaccurate results may occur because the linear weighting scheme will redistribute voxel variances differently than it redistributes voxel intensities.

The aim of this paper was to demonstrate this effect and to suggest a way of combining *t*-maps, *t*-thresholds of significance and contrast images to 'masked contrast images' that can be used for presentation and that may allow for a more precise localization of significant effects than *t*-maps alone. The suggested algorithm was applied to data from a previously published PET study with $[^{11}C]$Carfentanil that showed subcortical brain regions with increased μ-opiate receptor availability in abstinent alcoholics.

Materials and methods

Theory

Smoothing is an averaging process in which the intensity at a given voxel is replaced by a weighted average (i.e., a linear combination) of the values of voxels in some spatial neighborhood of that voxel. Common smoothing methods such as Gaussian filtering attribute the most weight to the (pretransformation value of) transformed voxel itself and are symmetric in the sense that the weights then decrease as a continuous function of spatial distance from the transformed voxel location, without any preferred direction (isotropic). A very simple example of such a scheme would consist of a bivariate statistic sampled from two subject groups (i.e., two variables per subject), but analyzed following a transformation such that for each subject, the transformed data associated with each variable is a linear combination of the original two variables. The weighted sum associated with one of the samples, say $\tilde{y}$ would be $\tilde{y} = w_1 y_1 + w_2 y_2$ where y_i and w_i are the original samples and weights ($i = 1, 2$) with standard deviations σ_1 and σ_2. If y_1 and y_2 are uncorrelated, the variance of $\tilde{y}$ is given by

$$\tilde{\sigma}^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2$$

A contrast, for example, a between-group difference $\Delta\mu$ can be tested by a *t*-statistic; the *t*-statistics on the *transformed* variable $\tilde{y}$, can be expressed in terms of contrasts $\Delta\mu_1$ and $\Delta\mu_2$ on each of the *original* variables y_1 and y_2:

$$\tilde{t} = const\frac{\Delta\tilde{\mu}}{\tilde{\sigma}} = const\frac{w_1\Delta\mu_1 + w_2\Delta\mu_2}{\sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2}}$$

where *const* depends only on sample size. On replacing $\Delta\mu_i$ with $t_i\sigma_i/const$, one obtains

$$\tilde{t} = \frac{w_1 t_1\sigma_1 + w_2 t_2\sigma_2}{\sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2}} \quad (1)$$

Assuming $t_1, t_2 > 0$, there is a ratio of weights w_1 and w_2 for which $\tilde{t}$ is maximal. As one can see directly from the vector notation below (equation (9)), this is when w_1 and w_2 correspond to 1/(relative variance) with relative variance being $\sigma_i^2/\Delta\mu_i$:

$$\frac{w_1}{w_2} = \frac{t_1/\sigma_1}{t_2/\sigma_2} \quad (2)$$

When the applied weights are proportional to 1/(relative variance), the maximal *t*-value will be associated with the variable having the smaller relative variance, and thus the ordering of the *t*-values may be reversed compared with their presmoothed values if the sample with the higher *t*-value is also the sample with the higher relative variance. The contrast itself is not similarly affected—if $\Delta\mu_1 > \Delta\mu_2$ and $w_1 > w_2$ then the same ordering is preserved in the transformed contrast.

We now seek the range of weights that will lead to a reversal of *t*-statistic ordering. Given the symmetry of the smoothing kernel and appropriate normalization of the weights (i.e., $w_1 + w_2 = 1$), $\tilde{y}_1$ and $\tilde{y}_2$ can be expressed as

$$\tilde{y}_1 = wy_1 + (1 - w)y_2$$

$$\tilde{y}_2 = (1 - w)y_1 + wy_2$$

and, if y_1 and y_2 are uncorrelated, the corresponding *t*-values $\tilde{t}_1$ and $\tilde{t}_1$ are given by equation (1). All weights are positive and as with Gaussian smoothing we assume greater weight in $\tilde{y}_1$ is given to y_1 than y_2, and vice versa for $\tilde{y}_2$, thus

$0.5 < w < 1$. As shown in Appendix A, the ordering of *t*-values will be reversed if $\sigma_1/t_1 > \sigma_2/t_2$ (thus $\sigma_1 > \sigma_2$) and

$$0.5 < w < 0.5 + \frac{1}{2}\sqrt{\frac{(\sigma_1/\sigma_2 - \sigma_2/\sigma_1) - (t_1/t_2 - t_2/t_1)}{(\sigma_1/\sigma_2 - \sigma_2/\sigma_1) + (t_1/t_2 - t_2/t_1)}} \quad (3)$$

Next, we extend the discussion to a setting in which the data is defined on a continuous spatial domain and smoothing has been performed. As commonly applied, Gaussian smoothing is linear, isotropic and stationary (independent on position), and so can be represented by convolution $\tilde{y}(x) = f(x) \otimes w(x)$ with $w(x)$ being the kernel function in the shape of a normal probability density function:

$$w(x) = \frac{\sqrt{\ln(2)/\pi}}{FWHM/2} e^{\frac{-x^2 \ln(2)}{(FWHM/2)^2}} \quad (4)$$

where *FWHM* denotes the 'full-width at half-maximum' of the smoothing kernel.

As a very idealized model for Gaussian smoothed data with two gray matter regions that are small compared with the *FWHM* of the smoothing kernel and surrounded by white matter with negligible absolute variability, let y_1 and y_2 be two point sources, separated by a distance d. With d_1 being the distance from source one to a given voxel on the line segment between source one and source two, the distance of this voxel to source two is $d_2 = d - d_1$ and the weights w_1, w_2 attributed to source one and source two are given by equation (4). Together with equation (2), one obtains the position of the maximum *t*-value:

$$d_1 = \frac{d}{2} - \frac{(FWHM/2)^2}{2d} \log_2\left(\frac{t_1\sigma_2}{t_2\sigma_1}\right) \quad (5)$$

For a small *FWHM*, this is the middle of both point sources. With increasing *FWHM*, the maximum shifts towards the point source with the lower relative variance.

For imaging data, which can be thought of as a discrete lattice representation of a variable defined on an underlying spatially continuous domain, we expand equation (1) to account for n samples $y_1, y_2 \ldots y_n$ that contribute to a smoothed voxel $\tilde{y}(x)$ (x = three-dimensional voxel coordinate) with the weights $w_1(x), w_2(x), \ldots, w_n(x)$:

$$\tilde{t}(x) = \frac{w_1(x)t_1\sigma_1 + \ldots + w_n(x)t_n\sigma_n}{\sqrt{[w_1(x)\sigma_1]^2 + \ldots + [w_n(x)\sigma_n]^2}} \quad (6)$$

The assumption for this equation is that all original samples y_i are uncorrelated, which, of course, is not the case for measured voxels. We therefore use y_i to represent independent *components* of an image rather than voxels. For example, a homogeneous region i with the spatial extent being represented as a voxel mask $m_i(x)$ enters equation (6) as one single sample y_i, and the weight $w_i(x)$ with which it contributes to a smoothed voxel $\tilde{y}(x)$ can be obtained from a convolution of the voxel mask with the smoothing kernel $w_s(x)$:

$$w_i(x) = m_i(x) \otimes w_s(x) \quad (7)$$

Note that the m_i may overlap, corresponding to different regions sharing some component of signal; that is, a voxel x may be a part of one homogenous set of voxels with respect to some component of the signal, and a different set with respect to another component.

The *t*-value of a smoothed voxel can also be expressed by using the vector notation for equation (6):

$$\tilde{t}(x) = \vec{t} \cdot \frac{\vec{S}(x)}{||\vec{S}(x)||} \quad (8)$$

where $\vec{t}$ is the $n \times 1$ vector of the original *t*-values and $\vec{S}(x)$ the $n \times 1$ vector of the weighted standard deviations $w_1(x)\sigma_1, w_2(x)\sigma_2, \ldots, w_n(x)\sigma_n$. As this is an inner product and the length of $\vec{S}(x)/||\vec{S}(x)||$ is one, one obtains:

$$\tilde{t}(x) = ||\vec{t}|| \cos(\phi(x)) \quad (9)$$

where $\phi(x)$ is the angle between $\vec{t}$ and $\vec{S}(x)$. The maximal possible *t*-value $||\vec{t}|| = \sqrt{t_1^2 + t_2^2 + \ldots + t_n^2}$ is obtained when $\vec{t}$ and $\vec{S}(x)$ are parallel, that is, when the original samples are weighted with 1/(relative variance). In three-dimensional space and for more than three point sources, this will not, in general, be an obtainable condition and the observed maximum T_{max} will be below the theoretical maximum $||\vec{t}||$. Still, the local maxima of $\tilde{t}(x)$ correspond to local minima of $|\phi(x)|$.

For reasons of simplicity, we do not distinguish between the point spread function $w_{PSF}(x)$ of the scanner and an additional smoothing kernel applied during image post-processing $w_{SPM}(x)$, since, in a typical voxelwise analysis, the former is much smaller and its contribution is negligible. However, one can easily account for both steps of smoothing by defining $w_s(x)$ as the combined kernel $w_s(x) = w_{PSF}(x) \otimes w_{SPM}(x)$. Assuming $w_{PSF}(x)$ to have the shape of a bell-curve (as $w_{SPM}(x)$), the full-width half-maximum of $w_s(x)$ is given by

$$FWHM_{total} = \sqrt{FWHM_{PSF}^2 + FWHM_{SPM}^2}$$

One-dimensional Simulations

All simulations in this paper were calculated with matlab (Mathworks, Natick, MA, USA). To allow for Monte Carlo simulations with small sample sizes, we did not use equation (6) in the one-dimensional simulation, since equation (6) refers to the true population contrast and standard deviation $\Delta\mu$ and σ and not their estimates from small samples (here, the latter are denoted $\Delta\hat{\mu}$ and $\hat{\sigma}$). Instead, we calculated a set of unsmoothed one-dimensional 'images' $f(x)$ and obtained the smoothed profiles $\tilde{y}(x)$ (each of which is meant to represent smoothed data from one subject) by convolution. A two sample *t*-test was then calculated from smoothed pixels.

Assuming uncorrelated Gaussian noise $\varepsilon_{pix}(x)$ in each pixel (measurement error) and interindividual variability ε_{reg_i} in each region, $f(x)$ was calculated for two groups of 'subjects' ($N = 2*1000$) from

$$f(x) = \mu_i + \varepsilon_{reg_i} + \varepsilon_{pix}(x)$$

The corresponding group difference $\Delta\mu(x)$ and the standard deviation $\sigma_{total}(x) = \sqrt{\sigma_{reg}^2 + \sigma_{pix}^2(x)}$ for the first

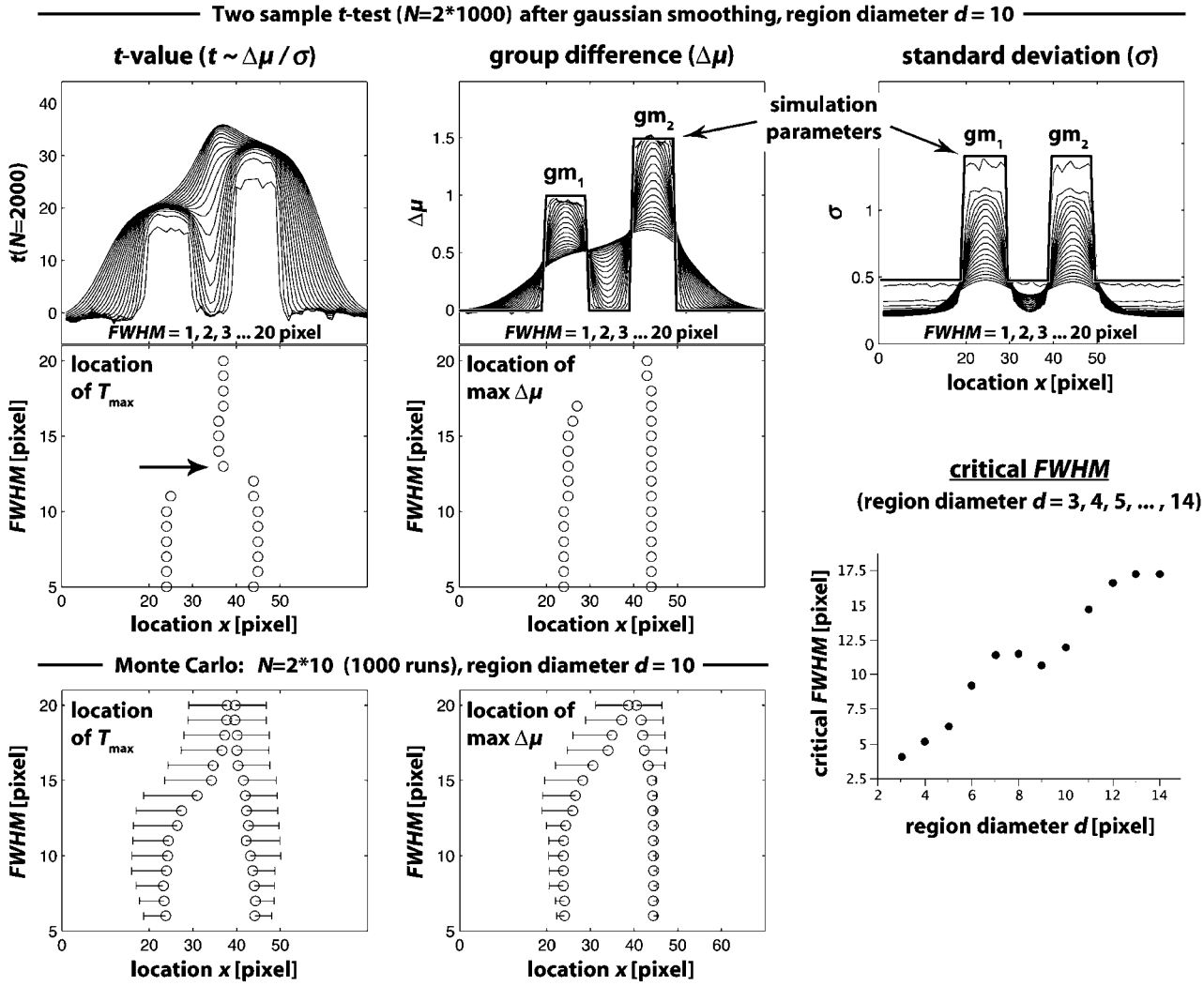

Figure 1 Effect of one-dimensional smoothing on $t(x)$, $\Delta\hat{\mu}(x)$ and $\hat{\sigma}(x)$ in a pixelwise two sample *t*-test. Left column: in most pixels, the *t*-value was increasing with *FWHM*. The highest increase was found in the white matter between gm_1 and gm_2. For *FWHM* ≥ 13 pixel, the *t*-value there exceeded the *t*-value in gray matter (arrow). Middle column: The location of the maximum group difference was less affected by spatial smoothing than that of T_{max} and no additional maximum between gm_1 and gm_2 occurs. Lower right: 'critical *FWHM*' (the *FWHM* at which a white matter T_{max} occurs) as a function of the region diameter *d* (*d* = width of gm_1, gm_2 and white matter in between).

simulation are shown in Figure 1 ('simulation parameters'). Width and distance of both gray matter regions was $d = 10$ pixel. For the gray matter regions (gm_1 and gm_2), we assumed a group difference $\Delta\mu_{gm1} = 1$ and $\Delta\mu_{gm2} = 1.5$. Normally distributed random values for $\varepsilon_{pix}(x)$ and ε_{reg_i} were calculated with matlab's function randn() and a standard deviation $\sigma_{reg} = 1$ and $\sigma_{pix} = 1$ (thus $\sigma_{total} = 1.41$) in both gray matter regions. In white matter, no group difference and a lower interindividual variability $\sigma_{reg} = 0.2$ was assumed. The measurement error in white matter was set to $\sigma_{pix} = \sqrt{0.2}$ (thus $\sigma_{total} = 0.49$) in accordance with the fact that a lower count rate in PET leads to a lower absolute, but a higher relative statistical error. After Gaussian smoothing with different full-width half-maximum (*FWHM* = 1, 2, …, 20 pixel), a *t*-test was calculated for each pixel, and the local maxima of $t(x)$ and $\Delta\hat{\mu}(x)$ (T_{max} and $\Delta\hat{\mu}_{max}$) were determined for each *FWHM* by a search algorithm starting in the middle of each gray matter region. For the same setting but a smaller number of subjects ($N = 2*10$), mean and standard deviation of the position of T_{max} and $\Delta\hat{\mu}_{max}$ were assessed by Monte Carlo simulation. We also calculated $f(x)$ with different region diameters ($d = 3, 4, 5, \ldots, 14$ pixel) and obtained the position of T_{max} from equation (9). For each *d* we increased the *FWHM* by steps of 0.1 until T_{max} shifted into white matter to obtain the 'critical *FWHM*'. It should be mentioned that a different number of pixels per region results in a different relative contribution from ε_{reg} and ε_{pix}.

To further illustrate the artifacts in Figure 1, we chose some modified profiles, including the (idealized) assumption of $\sigma = 0$ in the white matter between gm_1 and gm_2. The corresponding $\Delta\mu(x)$ and $\sigma_{total}(x)$ are shown in Figure 2. Unless mentioned explicitly, the parameters σ_{reg} and σ_{pix} were the same as in Figure 1.

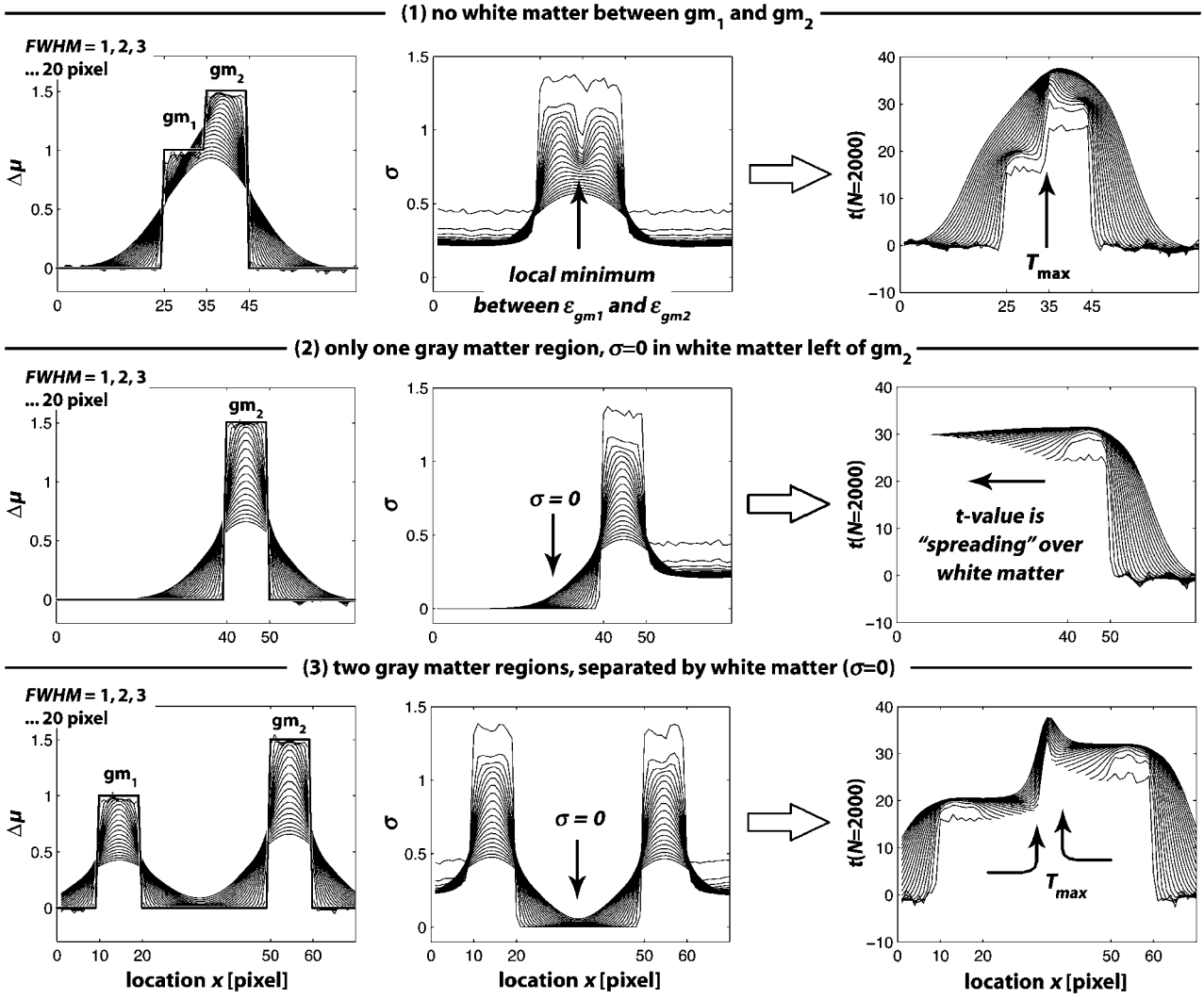

Figure 2 Effects from white matter. The simulation from Figure 1 was modified to illustrate the effects from white matter. Top row: without white matter, smoothing leads to a local minimum of $\hat{\sigma}(x)$ at the border between gm_1 and gm_2, and to a corresponding peak in $t(x)$ (black arrow). Middle row: Noise and variability in the white matter on the left side from gm_2 was set to zero. After smoothing, the *t*-value in gm_2 is 'spreading' to the left. Bottom row: white matter ($\sigma = 0$) allows gm_1 and gm_2 to interfere in spite of a large distance (30 pixel) and create an artificial T_{max}.

Two-dimensional Simulation

We used a two-dimensional model of three square gray matter regions (5×5 pixel each, Figure 3) and surrounding white matter with the same parameters σ_{reg} (between-subject variability) and σ_{pix} (measurement error) as in Figure 1. The group difference $\Delta\mu$ was zero in the gray matter region in the middle and was $\Delta\mu = 1$ in the other two regions. With $FWHM = 14$ pixel, weights $w_i(x)$ were calculated from equation (7) for all 'anatomic' regions (three gray matter regions, one white matter region, standard deviation σ_{reg}) and additionally for each pixel (standard deviation σ_{pix}). The weighted standard deviations $\vec{S}(x)$ were calculated and $\tilde{t}(x)$ was obtained from equation (9). $\tilde{t}$-isocontours were plotted between 95 and 100% of the observed maximum. For aesthetical reasons, we did not use the same resolution for region mask and smoothed image space, instead, we calculated $w_i(x)$ and $\tilde{t}(x)$ on a much finer grid.

[^{11}C]Carfentanil-PET

Measured PET data presented in this paper originate from a previously published PET study in abstinent alcoholics (Heinz *et al*, 2005). In 20 alcoholics, abstinent for 2 to 3 weeks, and in nine healthy control subjects, radioactivity distribution in the brain was measured with a GE Advance PET-scanner 0 to 66 min after intravenous injection of 700 MBq [^{11}C]Carfentanil, a highly selective μ-opiate receptor ligand. Stereotactically normalized parametric images of receptor availability (V_3'') were calculated with Logan's linearization and the occipital cortex as a reference region with negligible specific binding. For details, see Heinz *et al* (2005).

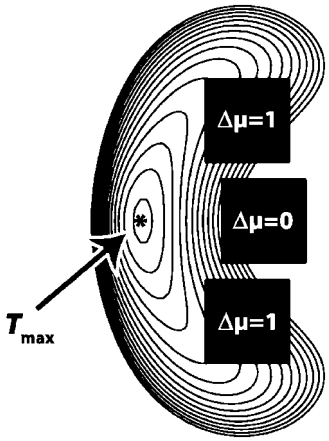

Figure 3 Two-dimensional simulation (worst case): *t*-isocontours from three gray matter regions (5×5 mm $= 5 \times 5$ pixel) and surrounding white matter, smoothed with *FWHM* = 14 mm. All gray matter regions had the same variance, yet only the upper and the lower region had a positive contrast ($\Delta\mu = 1$). Note that the nearest gray matter of T_{max} is the region with $\Delta\mu = 0$.

SPM Analysis

With SPM2, V_3'' images were smoothed with a 12-mm Gaussian kernel, and a voxelwise two-sample *t*-test was calculated. Unlike previously published (Heinz *et al*, 2005), we did not mask out white matter. We applied a voxel-level threshold of $P = 0.001$ (uncorrected) and confirmed that both striatal suprathreshold clusters survived correction for multiple comparisons with SPM's small volume correction and a mask for the striatal volume of interest (12.1 cm^3). The analysis was repeated with a voxel-level threshold of $P = 0.05$ (uncorrected) to obtain the corresponding *t*-threshold for later use.

Masking Algorithm for Contrast Images

To use contrast images for presentation, just as *t*-maps with significant clusters being displayed over corresponding anatomic sections, a masking algorithm is needed. Instead of using the original *t*-isocontours, which are affected by the investigated artifacts, we used a corresponding contrast-threshold that we had calculated separately for each cluster.

The following data were available from the SPM analysis:

- *t*-maps (file: *spmT_0002.img*).
- Contrast images (file: *con_0002.img*), containing the smoothed $\Delta V_3''$.
- A list of clusters (coordinates of T_{max}) that survived SPM's correction for multiple testing.
- *t*-thresholds for $P = 0.001$ and $P = 0.05$.

The mask that was applied to SPM's contrast images was calculated as illustrated in Figure 4. For each significant region, we first obtained the original SPM-cluster from *t*-maps with a region growing algorithm that starts at T_{max} and includes all voxels with a *t*-value above the original threshold (here $P = 0.001$). The mean contrast in this cluster was subsequently used for thresholding and all contiguous voxels that met both the $\Delta V_3''$ threshold and satisfied $P < 0.05$ (to ensure some statistical evidence) were included in the resulting region.

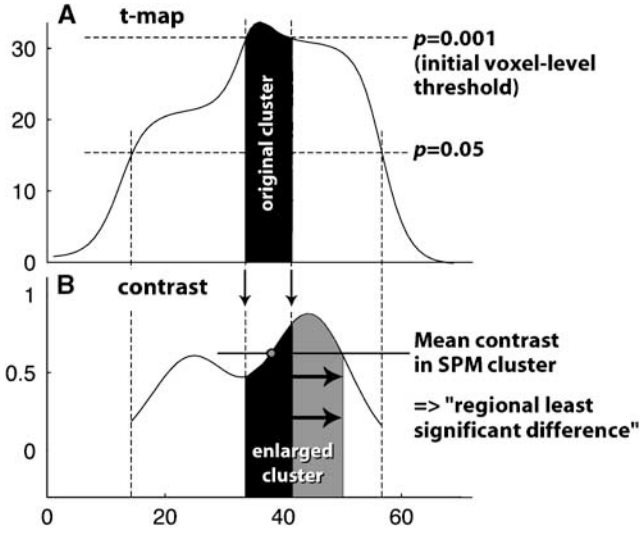

Figure 4 Suggested algorithm to create masked contrast images. (**A**) For each significant cluster, a mask (black bar) is generated from *t*-maps and the initial voxel-level *t*-threshold ($P = 0.001$) and subsequently applied to contrast images (**B**). The mean contrast in this cluster is calculated, and the cluster is enlarged by including adjacent voxels (grey area) that exceed this mean contrast. Voxels that do not reach the $P = 0.05$ level are not included. The depicted profiles were taken from Figure 1 (*FWHM* = 14 pixel).

Results

Simulations

The effect of Gaussian smoothing on $t(x)$, $\Delta\hat{\mu}(x)$ and $\Delta\hat{\sigma}(x)$ in a pixelwise two sample *t*-test is shown in Figure 1. In most pixels, the *t*-value was strictly increasing with *FWHM* (for *FWHM* = 3, 4, 5, …, 20). Unlike the shape of $\Delta\hat{\mu}(x)$, which changed with *FWHM* in a way one might guess intuitively, $t(x)$ showed an irregular pattern with the highest increase in the white matter between gm_1 and gm_2 (Figure 1, top left). Slightly decreasing *t*-values were observed only at higher *FWHM* (≥ 10 pixel) at the border between gray and white matter. At *FWHM* > 12 pixel, there was only one single T_{max} in the middle of gm_1 and gm_2 (Figure 1, arrow), which slightly shifted back towards gm_2 when *FWHM* was further increased. The 'critical *FWHM*' at which such a white matter peak occured was lower than 150% of the region diameter *d* in all simulations ($d = 1, 2, 3, \ldots, 14$). The position of $\Delta\hat{\mu}_{max}$ was much more stable: the local maximum in gm_1 was preserved until *FWHM* = 18 pixel, the maximum in gm_2 was located within the gray matter borders until *FWHM* = 28 pixel (not shown). The higher

robustness of $\Delta\hat{\mu}_{max}$ as opposed to T_{max} was also obvious in the Monte Carlo simulation (Figure 1, bottom): a higher variability of the position of T_{max} than that of $\Delta\hat{\mu}_{max}$ was observed for all simulated *FWHM*. At *FWHM* = 6 pixel, T_{max} of gm_1 was already outside the boundaries of gm_1 in 8% of all realizations.

Results from three modified profiles are depicted in Figure 2:

- Without white matter between gm_1 and gm_2 (Figure 2, top row), smoothing resulted in a local minimum in $\hat{\sigma}(x)$ at the border between gm_1 and gm_2 and a corresponding peak in $t(x)$. When ε_{reg} was set to zero (but not ε_{pix}), this peak disappeared (not depicted).
- In white matter with a hypothetical variance $\sigma_{total} = 0$ (Figure 2, middle row) and a $\Delta\mu = 0$, the standard deviation and $\Delta\hat{\mu}$ after smoothing solely reflect the contribution from the adjacent gm_2. Accordingly, in Figure 2, the *t*-value left of gm_2 was close to that within gm_2 (theoretically, if gm_2 was only one pixel wide and if the contribution from white matter on the right side of gm_2 was negligible, the *t*-value would be exactly the same). The horizontal arrow suggests looking at this effect as if the *t*-value was 'spreading' over white matter with increasing *FWHM* (no *t*-values were plotted for a denominator $\hat{\sigma}(x) < 0.001$).
- When white matter with $\sigma_{total} = 0$ was surrounded by two gray matter regions (Figure 2, bottom row), smoothing resulted in a peak $t(x)$ in the middle of gm_1 and gm_2, which first occurred when both 'spreading' *t*-values met exactly in the middle. For higher *FWHM*, T_{max} shifted towards gm_2 (the region with the lower relative variance). Interestingly—unlike in the simulation without white matter (Figure 2, top row)—interindividual variability ε_{reg} (in addition to the measurement error ε_{pix}) was not needed for this effect.

In the two-dimensional simulation (Figure 3), the maximum *t*-value was shifted to the left by the gray matter region in the middle (with $\Delta\mu = 0$). Note that in the chosen setting, T_{max} was still closer to the gray matter region in the middle than to the regions with a positive contrast.

[^{11}C]Carfentanil PET

SPM analysis of [^{11}C]Carfentanil PET-data confirmed significantly increased μ-opiate receptor availability in the bilateral ventral striatum. Without masking out white matter, the maximum *t*-value in the left hemisphere was found between ventral striatum and frontal cortex (Figure 5, top left), close to a local *minimum* of the unsmoothed $\Delta V_3''$. Masked contrast images (Figure 5, bottom) were more symmetrical, including more voxels in the middle of the left ventral striatum where the maximum $\Delta V_3''$ was found.

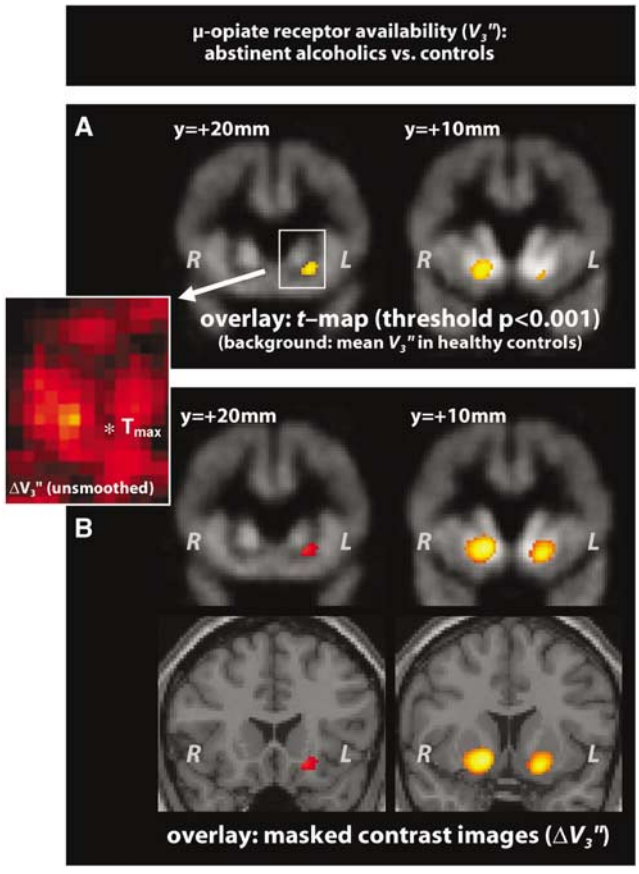

Figure 5 Results from a PET-study with [^{11}C]Carfentanil. μ-opiate-receptor availability V_3'' in abstinent alcoholics was compared with that in healthy controls. (**A**) Traditionally thresholded *t*-maps show significantly elevated V_3'' in the bilateral ventral striatum. However, the *maximum* *t*-value of the left cluster was found in adjacent white matter ($y = +20$ mm), near the local *minimum* of unsmoothed $\Delta V_3''$ (see inset, T_{max} = white star '*'). (**B**) SPM's contrast images (smoothed $\Delta V_3''$), masked as described, show a more symmetrical pattern with a T_{max} within the borders of the bilateral ventral striatum ($y = +10$ mm).

In the right striatum, group differences were stronger and the standard deviation showed less local variation. Accordingly, masked contrast images were very similar to the original *t*-map.

Discussion

In addition to the loss of resolution that is inherent in spatial smoothing, the nonlinear interactions with the voxel variances (which affect *t*-maps, but not contrast images) can result in sometimes-counterintuitive artifacts, as demonstrated in this paper. These effects include a displacement of T_{max} *away* from a region with a higher relative variance, and a displacement of T_{max} towards the middle of two or more regions with a positive contrast and independent error terms. In the latter case, a displaced peak differs from a smooth, combined

peak in contrast images in that it manifests at much lower levels of smoothing, and is spatially narrowed in a manner such that it may be mistakenly identified as an additional finding. The position of this peak depends on the relative variances of the surrounding regions so that the 'nearest gray matter' algorithm may not always determine the correct anatomic region.

This can be understood by looking at two effects:

- smoothing causes the independent error terms of two or more neighboring gray matter regions to be averaged, resulting in a minimum variance between these regions
- white matter with little statistical noise and low absolute interindividual variability 'inherits' the *t*-value of adjacent gray matter, and peaks in *t*-maps may become broader than the corresponding peaks in contrast images. This not only leads to a higher uncertainty of the position of T_{max} (as demonstrated in the Monte Carlo simulation), white matter also may serve as a 'bridge' between two or more interacting gray matter regions.

In our one-dimensional simulation, Gaussian smoothing with an *FWHM* in the same order of magnitude as the distance of the involved regions was already critical. Depending on the actual structure, this effect may even be stronger after three-dimensional smoothing. While such a *t*-shift is less likely to affect functional mapping of the cortex, it may seriously impair localization of subcortical findings, as for example, in the presented PET study with [^{11}C]Carfentanil, where a maximum *t*-value was observed in the white matter between ventral striatum (where higher μ-receptor availability was expected) and the frontal cortex (where alcoholics also displayed a higher μ-receptor availability, though not significant). Smaller misplacements may just be an aesthetical concern or may raise the question of how to present the results in a convincing way. Larger misplacements may get masked out (white matter mask) at the expense of sensitivity. Findings may also go undetected when the T_{max} shifts into a neighboring gray matter region with lower relative variance. Finally, misinterpretations may occur in brain areas with more than two gray matter regions, as in our two-dimensional (worst case) simulation, where the 'nearest gray matter' algorithm would not have detected the two regions that contributed to T_{max}. Likewise, one has to be careful when assigning peaks to functional subunits, for example, of the thalamus or the ventral striatum.

In the literature, different smoothing kernels are usually compared with regard to the *sensitivity* of detecting significant regions (Hopfinger *et al*, 2000). Using a spatially stationary noise model, Worsley *et al* (1996) also describe the effect on the *local* characteristic of the *t*-map, notably a broadened peak between the two original maxima. However, the most striking effects, including a narrowly focused but shifted peak as in our simulations, only occur when estimating a *local* variance (Friston *et al*, 1991) instead of a pooled variance across voxels (Worsley *et al*, 1992), and when smoothing across different types of tissue. To our knowledge, these effects have not yet been investigated in the literature.

Several methods have been suggested to improve spatial precision of statistical maps. Clearly, reanalyzing the data with a smaller smoothing kernel (Worsley *et al*, 1996) would have reduced the described artifacts, but at the expense of statistical significance. Another approach is to incorporate spatial *a priori* assumptions. Davatzikos *et al* (2001) have described an atlas-based adaptive smoothing to avoid smoothing across anatomic boundaries, which theoretically prevents the described artifacts. However, including *a priori* information either increases the complexity of the analysis or limits the advantages of a voxel-based (no *a priori* assumptions) over traditional ROI analysis. Accordingly, 'classical statistical maps' (Penny and Friston, 2004) are still the most popular approach for explorative imaging.

While all established methods have in common that the reported values are statistical measures related to a *level of confidence* by which a 'true effect' was observed, the method suggested in this paper follows a different rationale. We do rely on established methods of statistical mapping (in particular those implemented in SPM2, however, other methods as a starting point are also possible) to *detect* significant regions, but once a cluster is considered significant, we suggest attributing voxel-wise color-codes to the estimated parameter itself instead of the level of confidence by which it is greater than zero.

The limitations of the presented method are those of any voxel-based analysis. They are powerful tools for explorative imaging, but they are not necessarily a replacement for traditional region-of-interest (ROI) analysis when a regional *a priori* hypothesis exists and when an ROI definition is anatomically and physiologically justified. Comparing voxel-based with ROI-based analyses is not subject of this paper, but it should be mentioned that the proposed method was primarily developed for assessment of *subcortical* regions for which an ROI analysis may indeed be a more straightforward approach. In voxel-based analyses, the type I error must be reduced by using a rather conservative voxel-level threshold (e.g., $P<0.001$), which may result in a loss of statistical power. Applying such a single *t*-threshold as a first step has shown to be a simple and powerful method to *detect* significant clusters (e.g., by considering the cluster size). However, such a *t*-threshold is not necessarily the best choice to determine the *outline* of a neurobiologic effect, which leaves room for other methods of masking such as the one presented in this paper. Using the *contrast* for color-coding, it may seem natural to use it also to define the outline of a displayed region.

Indeed, we have shown that this may provide further information about the locus of, for example, a group difference, albeit in an exploratory form outside the framework of hypothesis testing. It should be noted that our masked contrast images differ from the original thresholded *t*-maps only if there is a local variation in the variance of the smoothed voxels. The suggested method can therefore be understood as an attempt to stick closely with the current standards while removing the described smoothing artifacts.

We believe that adding a further step to the analysis by going from thresholded *t*-maps to masked contrast images does not substantially increase the overall complexity of the analysis. On the contrary, it may give the investigator a more direct sense of the underlying data and may improve spatial precision when a significant region occurs. The simplicity of the suggested method and the fact that we do not make further assumptions make it a useful complement for established methods for statistical mapping.

References

Ashburner J, Friston KJ (1999) Nonlinear spatial normalization using basis functions. *Hum Brain Mapping* 7:254–66

Ashburner J, Friston KJ, Penni W (eds) (2003) *Human Brain Function, 2nd ed, Chapt. Part II—Imaging Neuroscience—Theory and Analysis*. Amsterdam: Elsevier

Davatzikos C, Li HH, Herskovits E, Resnick SM (2001) Accuracy and sensitivity of detection of activation foci in the brain via statistical parametric mapping: a study using a PET simulator. *Neuroimage* 13:176–84

Friston KJ, Frith CD, Liddle PF, Frackowiak RSJ (1991) Comparing functional (PET) images: the assessment of significant change. *J Cereb Blood Flow Metab* 11:690–9

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ (1995) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapping* 2:189–210

Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC (1994) Assessing the significance of focal activations using their spatial extent. *Hum Brain Mapping* 1:214–20

Heinz A, Reimold M, Wrase J, Hermann D, Croissant B, Mundle G, Dohmen BM, Braus DH, Schumann G, Machulla H-J, Bares R, Mann K (2005) Stable elevations in striatal μ-opioid receptor availability in detoxified alcoholics correlate with alcohol craving a [^{11}C]Carfentanil PET study. *Arch Gen Psychol* 62:57–64

Hopfinger JB, Büchel C, Holmes AP, Friston KJ (2000) A study of analysis parameters that influence the sensitivity of event-related fMRI analyses. *Neuroimage* 11: 326–33

Penny W, Friston K (2004) Classical and Bayesian inference in fMRI. In: *Advanced Image Processing in Magnetic Resonance Imaging* (Landini L, ed) New York: Marcel Dekker

Rosenfeld A, Kak AC (1982) *Digital Picture Processing*. New York: Academic Press

Worsley KJ, Evans AC, Marret S, Neelin PA (1992) Three-dimensional statistical analysis for CBF activation studies in human brain. *J Cereb Blood Flow Metab* 12:900–1180

Worsley KJ, Marrett S, Neelin P, Evans AC (1996) Searching scale space for activation in PET images. *Hum Brain Mapping* 4:74–90

Appendix A

We seek the values of w, $0.5 < w < 1$, such that $\tilde{t}_2 > \tilde{t}_1$ while $t_1 > t_2$, or $\mathrm{DIF}(w) = \tilde{t}_1 - \tilde{t}_2 < 0$.

Note that $\mathrm{DIF}(0.5) = 0$, and $\mathrm{DIF}(1) = t_1 - t_2 > 0$, and DIF is a differentiable function of w. The derivative, evaluated at $w = 0.5$ is

$$\frac{4\sigma_1\sigma_2(\sigma_2 t_1 - \sigma_1 t_2)}{(\sigma_1^2 + \sigma_2^2)^{3/2}}$$

and this will be negative (i.e., $\mathrm{DIF}(w)$ will be negative immediately to the right of $w = 0.5$) when $\sigma_1/t_1 > \sigma_2/t_2$.

Therefore, if this condition pertains and $\mathrm{DIF}(w)$ has exactly one root in the open interval $(0.5, 1)$, it must be the case that DIF is negative between 0.5 and the root, and positive between the root and 1, and the ordering of the *t*-map will be reversed compared with the original values on the interval $0.5 < w < root$.

We therefore seek to show that $\mathrm{DIF}(w)$ has exactly one root between 0.5 and 1. $\mathrm{DIF} = 0$ is equivalent to $\tilde{t}_1^2 = \tilde{t}_2^2$ (for $\tilde{t} > 0$), which can be written as

$$\frac{a^2w^2 + 2abw + b^2}{cw^2 - 2dw + d} = \frac{a^2w^2 - 2a(a+b)w + (a+b)^2}{cw^2 + 2(d-c)w - (d-c)} \tag{A1}$$

with

$$\begin{aligned} a &= t_1\sigma_1 - t_2\sigma_2 \\ b &= t_2\sigma_2 \\ c &= \sigma_1^2 + \sigma_2^2 \\ d &= \sigma_2^2 \end{aligned}$$

This leads to a cubic equation in w (the 4th power term cancels). Noting that 0.5 is a solution of equation (A1), the remaining two solutions of equation (A1) are the roots of the quadratic

$$\begin{aligned} &(a^2d + abc)w(w-1) \\ &+ [b^2(d - c/2)ad(a/2 + b)] \end{aligned} \tag{A2}$$

These roots, in terms of the original variables, are

$$0.5 \times \left[1 \pm \sqrt{\frac{(\sigma_1/\sigma_2 - \sigma_2/\sigma_1) - (t_1/t_2 - t_2/t_1)}{(\sigma_1/\sigma_2 - \sigma_2/\sigma_1) + (t_1/t_2 - t_2/t_1)}}\right] \tag{A3}$$

Finally, given the conditions on t_1, t_2, $\tilde{t}_1$ and $\tilde{t}_2$, the expression under the radical sign must be between 0 and 1, and therefore the larger of the 2 roots falls between 0.5 and 1, and is the only solution to equation (A1) in that interval. Values of w falling between 0.5 and the larger root of equation (A1) will lead to a reversal of the ordering of the *t*-values.