



Commentary: Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates

Karsten Mueller^{1*}, Jöran Lepsien¹, Harald E. Möller¹ and Gabriele Lohmann^{2,3}

¹ Nuclear Magnetic Resonance Unit, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany,

² Department of Biomedical Magnetic Resonance, University Hospital Tuebingen, Tuebingen, Germany, ³ Magnetic Resonance Center, Max Planck Institute for Biological Cybernetics, Tuebingen, Tuebingen, Germany

Keywords: fMRI, functional magnetic resonance imaging, false positive results, false positive error, random field theory, family-wise error rate

A commentary on

Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates

by Eklund, A., Nichols, T. E., and Knutsson, H. (2016). *Proc. Natl. Acad. Sci. U.S.A.* 113, 7900–7905. doi: 10.1073/pnas.1602413113

OPEN ACCESS

Edited by:

Vladimir Litvak,
University College London Institute of
Neurology, United Kingdom

Reviewed by:

Gerard R Ridgway,
University of Oxford, United Kingdom
Anders Eklund,
Linköping University, Sweden

*Correspondence:

Karsten Mueller
karstenm@cbs.mpg.de

Received: 27 April 2017

Accepted: 14 June 2017

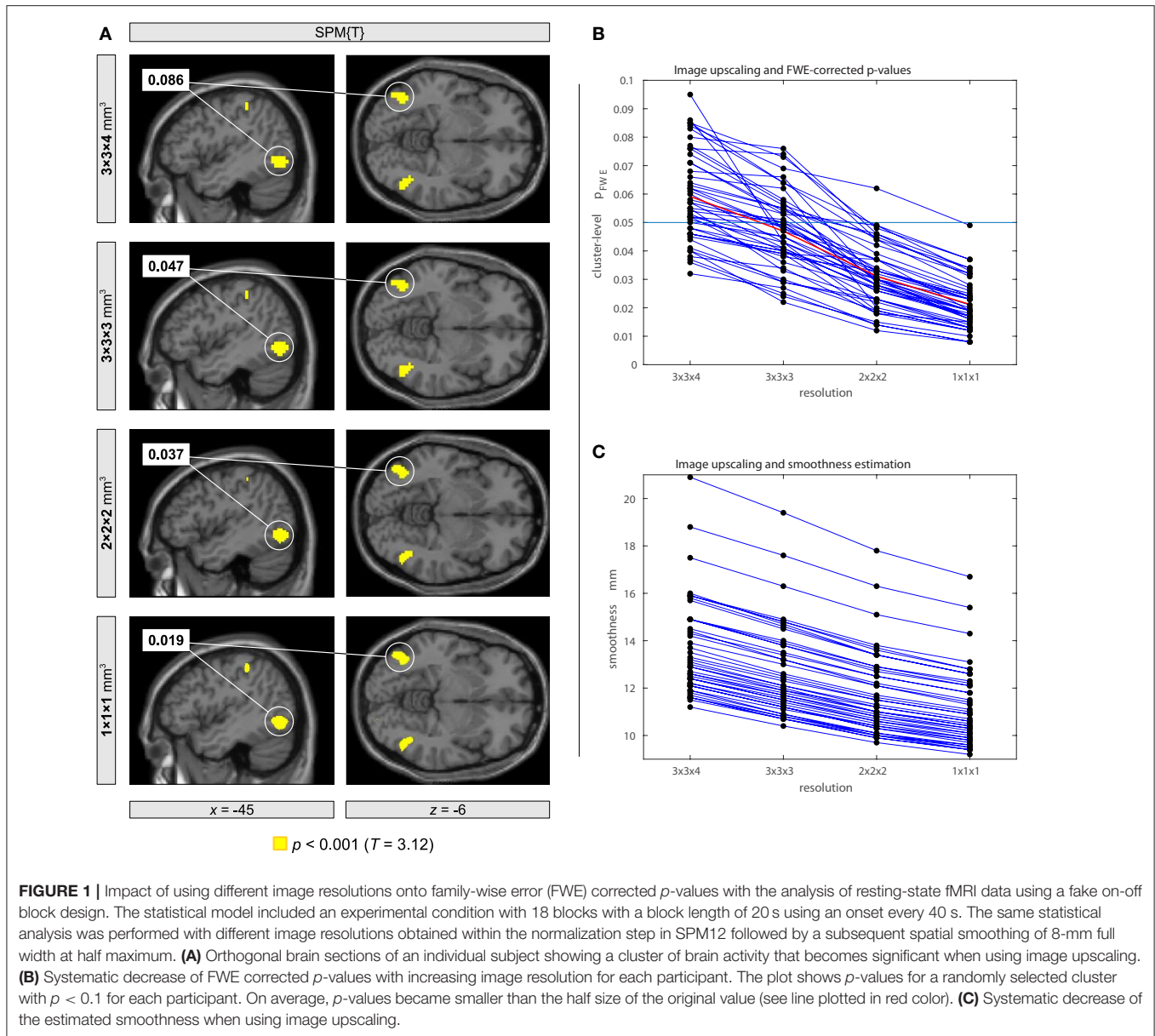
Published: 28 June 2017

Citation:

Mueller K, Lepsien J, Möller HE and
Lohmann G (2017) Commentary:
Cluster failure: Why fMRI inferences
for spatial extent have inflated
false-positive rates.
Front. Hum. Neurosci. 11:345.
doi: 10.3389/fnhum.2017.00345

In a recent manuscript, Eklund et al. (2016) reported inflated false positive rates in functional MRI (fMRI) using several common software packages. Here we would like to draw attention to an important aspect that was not addressed in this publication. Specifically, we would like to note that statistical inferences obtained using the random field theory depend heavily on a preprocessing parameter not discussed by Eklund et al. (2016), namely the spatial resolution to which the data sets are resampled and interpolated during pre-processing. This resampling is needed to align the data to a common anatomical template that is essential to perform group analyses (also often called *normalization*). Eklund et al. (2016) used the default setting of $2 \times 2 \times 2$ mm³. In response to Eklund's paper, Flandin and Friston (2016) used a different setting of this parameter, namely $3 \times 3 \times 3$ mm³. Together with a more stringent initial cluster-forming threshold, they did not observe inflated false positive rates. However, the $2 \times 2 \times 2$ mm² setting is the default in two major software packages (SPM, FSL), and in previous work, Friston and colleagues Hopfinger et al. (2000) stated that resampling to $2 \times 2 \times 2$ mm³ renders the analysis "more sensitive." In other words, at present it is unclear what a valid setting for this parameter should be. Therefore, we think that it is extremely relevant to assess its influence on statistical inference.

For this purpose, we analyzed a group of 47 resting-state fMRI data sets with a spatial resolution of $3 \times 3 \times 4$ mm³ and 300 volumes used in a preceding study Mueller et al. (2016). Using a strategy analogous to Eklund et al. (2016) we imposed various fake designs including block- and event-related types. We tested the following resampling parameters: $3 \times 3 \times 4$, $3 \times 3 \times 3$, $2 \times 2 \times 2$, and $1 \times 1 \times 1$ mm³. Using SPM12 with family-wise error (FWE) correction based on the random field theory, we first evaluated each data set separately. We found that with higher resampling resolutions, the FWE-corrected *p*-values decrease systematically so that more and more false positives occur. **Figure 1A** shows a typical result. We obtained a systematic effect in all of the 47 data sets (**Figure 1B**). We also observed a systematic effect of image upscaling onto smoothness estimation (**Figure 1C**). Furthermore, we performed a group-level inference in which all 47 data sets were pooled. Again, we observed that the FWE-corrected *p*-values decreased systematically with higher resampling resolutions.



In other words, it appears that there is a *systematic* dependence of the false positive rate on the resampling parameter with smaller voxel sizes leading to smaller FWE-corrected *p*-values and hence more false positives. While some dependence on preprocessing parameters may be inevitable, a systematic dependence of this type is clearly worrisome, because researchers may be tempted to interpolate their data until the desired statistical significance level is reached. Statistical inference should certainly not depend in such a systematic way on a preprocessing parameter that can be set *ad libitum*. Clearly, this issue requires further in-depth analysis.

AUTHOR CONTRIBUTIONS

KM communicated carried out the data analyses, and wrote the first draft of the letter. All authors assisted with the conceptual approach and contributed to the writing.

ACKNOWLEDGMENTS

We are thankful to Professor Anders Eklund for his very helpful communications.

REFERENCES

- Eklund, A., Nichols, T. E., and Knutsson, H. (2016). Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci. U.S.A.* 113, 7900–7905. doi: 10.1073/pnas.1602413113
- Flandin, G., and Friston, K. J. (2016). Analysis of family-wise error rates in statistical parametric mapping using random field theory. *arXiv* 1606.08199v1.
- Hopfinger, J. B., Büchel, C., Holmes, A. P., and Friston, K. J. (2000). A study of analysis parameters that influence the sensitivity of event-related fMRI analyses. *Neuroimage* 11, 326–333. doi: 10.1006/nimg.2000.0549
- Mueller, K., Arelin, K., Möller, H. E., Sacher, J., Kratzsch, J., Luck, T., et al. (2016). Serum BDNF correlates with connectivity in the (pre)motor hub in the aging

human brain—a resting-state fMRI pilot study. *Neurobiol. Aging* 38, 181–187. doi: 10.1016/j.neurobiolaging.2015.11.003

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Mueller, Lepsien, Möller and Lohmann. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.