*Letter to the Editor*

# Lack of Methodological Rigor for Task-Based Functional Magnetic Resonance Imaging: Injury-Related Fear or Failure to Correct?

Dear Editor:

We read with interest a recent report in the *Journal of Athletic Training*, "Neuroplasticity in Corticolimbic Brain Regions in Patients After Anterior Cruciate Ligament Reconstruction."[1] It is exciting to see neuroscience-based methods, specifically brain functional magnetic resonance imaging (fMRI), applied by sports medicine researchers to answer novel research questions. However, the methodological approaches used in the referenced manuscript[1] do not comply with contemporary standards of statistical analyses and reporting for fMRI studies,[2–5] such that the results are difficult to interpret. Our goal with this letter is to highlight the major analytical concerns and reinforce the concept that minimum analytic standards must be applied if task-based fMRI data are to inform and innovate sports medicine practice. Notably, the summarized concerns are not our unique recommendations but rather the analytical and reporting standards that have been established by experts in the neuroscience community for many years.[5]

## MULTIPLE-COMPARISONS CORRECTION AND STATISTICAL INFERENCES

To analyze task-based fMRI data, statistical maps are created to identify regions of the brain (ie, voxels) with increased activity in response to a manipulation or stimulus relative to a control or rest condition. A typical functional neuroimaging volume contains approximately 130 000 voxels (variation based on acquisition parameters), requiring thousands of statistical tests to contrast or determine voxels that demonstrate a significant response to a stimulus relative to rest or another condition. The sheer magnitude of statistical comparisons results in expected false-positives that require application of an activation threshold and multiple-comparisons correction to decipher task-related signal versus noise. Specifically, for task-based fMRI, voxels without activation above a statistical threshold are discarded (ensuring that the signal is task related beyond noise or the comparison condition), and the remaining voxels must survive a multiple-comparisons correction to minimize the degree of false-positives to predictable levels. Numerous ways of applying such thresholds and corrections are available for considering the unique data structure of fMRI (eg, cluster based, voxelwise and threshold-free cluster enhancement),[6,7] of which some have become the default settings in many fMRI statistical analysis packages.

The fMRI analysis in the manuscript in question did not provide any level of thresholding or multiple-comparisons correction. Use of an uncorrected approach in fMRI can result in a degree of false-positives so severe that 1 research group[8] published the infamous "dead salmon paper," in which a deceased salmon demonstrated "significant neural activity" when exposed to images and the completed analysis was uncorrected. However, with appropriate corrections applied, no significant signal was detected, as would be expected with deceased tissue.[8] This was a tongue-in-cheek report to emphasize the need for minimal statistical corrections and thresholding in fMRI analyses, highlighting that a portion of "significant" task results reported are in fact false-positive indicators of relative brain activations when the data are uncorrected. In other words, without applying these fundamental statistical controls, it is impossible to estimate the type I error rate, thus making any finding unreliable. The lack of thresholding and multiple-comparisons correction is so fundamentally flawed in fMRI analyses that neuroimaging journals often will not even consider a submission without these essential statistical corrections.[9]

## *P* HACKING, POST HOC REGION OF INTEREST SELECTION, AND CIRCULAR ANALYSES

The authors indicated that regions of interest (ROIs) were not determined a priori as typically recommended and instead were selected using a "qualitative post hoc" approach. The selection of ROIs after the primary analysis is referred to as *circularity* (or "double dipping"), which leads to vastly inflated effect sizes and is widely considered an unacceptable practice.[4,10,11] The inflation of findings is readily apparent in Table 2 as all 22 ROIs selected were different between groups, when the automated anatomical labeling approach resulted in 90 possible ROIs.[12] This mode of "cherry picking" or "self-selecting" ROIs in task-based fMRI is a neuroimaging version of *P* hacking, ie, examining the data before making ROI selections. Although inflation due to circularity has plagued numerous published studies, the detrimental consequences of such an inference are compounded with the combination of circular analyses of uncorrected and unthresholded data,[11] as completed in this recent referenced manuscript.[1]

## TREATMENT OF TASK CONDITIONS

The use of the picture imagination task, with depictions of sport-specific activities and activities of daily living (ADLs), to compare task-related activity between participants with anterior cruciate ligament reconstruction versus healthy participants is intriguing. However, the combination

of sport and ADL images is a puzzling data presentation. The use of ADLs as a visual control for sport images could be an elegant design to isolate sport-specific imagery and a potential fear response, but this between-conditions comparison does not appear to have been applied in the between-groups analysis. A secondary analysis comparing sport and ADL images was completed but only in the reconstruction group; thus, whether sport or ADL image processing is different between or within groups is unknown. Furthermore, by neither thresholding nor correcting for multiple comparisons, the authors' decision to compute an average blood oxygen level-dependent signal across an anatomically derived ROI average activity in both task-relevant and nonrelevant voxels is puzzling. Given the lack of identification of image-specific responsive voxels, it is not possible to determine the validity of the authors' suggestion of a similar neurologic fear response to images of sitting and reading a book to images of sport maneuvers.

## CONCLUSIONS

The purpose of this letter to the editor is to indicate that the methodological approach used in the recently published manuscript[1] did not achieve the accepted standards of statistical analysis for task-based fMRI measures. Readers should therefore be extremely cautious in drawing conclusions from the reported results. We encourage the authors to reanalyze their data based on these recommendations so that the findings are more interpretable and meaningful to the sports medicine community.

*Dustin R. Grooms, PhD, ATC*
Ohio University, Athens

*Alexis B. Slutsky-Ganesh, PhD*
Emory University School of Medicine, Atlanta, GA

*Janet E. Simon, PhD, ATC*
Ohio University, Athens

*Manish Anand, PhD*
*Gregory D. Myer, PhD*
Emory University School of Medicine, Atlanta, GA

*Jed A. Diekfuss, PhD*
Emory University School of Medicine, Atlanta, GA

## REFERENCES

1. Baez S, Andersen A, Andreatta R, Cormier M, Gribble PA, Hoch JM. Neuroplasticity in corticolimbic brain regions in patients after anterior cruciate ligament reconstruction. *J Athl Train*. 2021;56(4):418–426. doi:10.4085/JAT0042-20
2. Poldrack RA, Fletcher PC, Henson RN, Worsley KJ, Brett M, Nichols TE. Guidelines for reporting an fMRI study. *Neuroimage*. 2008;40(2):409–414. doi:10.1016/j.neuroimage.2007.11.048
3. Poldrack RA, Mumford JA. Independence in ROI analysis: where is the voodoo? *Soc Cogn Affect Neurosci*. 2009;4(2):208–213. doi:10.1093/scan/nsp011
4. Vul E, Pashler H. Voodoo and circularity errors. *Neuroimage*. 2012;62(2):945–948. doi:10.1016/j.neuroimage.2012.01.027
5. Nichols TE, Das S, Eickhoff SB, et al. Best practices in data analysis and sharing in neuroimaging using MRI. *Nat Neurosci*. 2017;20(3):299–303. doi:10.1038/nn.4500
6. Woo CW, Krishnan A, Wager TD. Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *Neuroimage*. 2014;91:412–419. doi:10.1016/j.neuroimage.2013.12.058
7. Nichols TE. Multiple testing corrections, nonparametric methods, and random field theory. *Neuroimage*. 2012;62(2):811–815. doi:10.1016/j.neuroimage.2012.04.014
8. Bennett CM, Baird AA, Miller MB, Wolford GL. Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: an argument for multiple comparisons correction. *Neuroimage*. 2009;47(suppl 1):S125.
9. Roiser JP, Linden DE, Gorno-Tempini ML, Moran RJ, Dickerson BC, Grafton ST. Minimum statistical standards for submissions to *Neuroimage: Clinical*. *Neuroimage Clin*. 2016;12:1045–1047. doi:10.1016/j.nicl.2016.08.002
10. Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci*. 2009;12(5):535–540. doi:10.1038/nn.2303
11. Kriegeskorte N, Lindquist MA, Nichols TE, Poldrack RA, Vul E. Everything you never wanted to know about circular analysis, but were afraid to ask. *J Cereb Blood Flow Metab*. 2010;30(9):1551–1557. doi:10.1038/jcbfm.2010.86
12. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 2002;15(1):273–289. doi:10.1006/nimg.2001.0978