

**Are your MRI contrast agents cost-effective?**

Learn more about generic Gadolinium-Based Contrast Agents.



**FRESENIUS  
KABI**

caring for life

**AJNR**

**Reply:**

S. Haller, K.-O. Lovblad, P. Giannakopoulos and D. Van De Ville

*AJNR Am J Neuroradiol* published online 18 July 2013

<http://www.ajnr.org/content/early/2013/07/18/ajnr.A3699.citation>

This information is current as  
of May 9, 2024.

## REPLY:

We thank Diciotti et al for the interest in our study and their thoughtful comments regarding the methodology that was used to obtain part of the results reported in our manuscript. We acknowledge that feature selection is an important part of learning algorithms and still the subject of research in high-dimensional (and structured) datasets. We had already mentioned this “peeking” issue raised by Diciotti et al as a limitation in our manuscript.

Motivated by the comment by Diciotti et al, we repeated our data analysis by using feature selection within the cross-validation folds. In particular, we reanalyzed the data by using Relief feature selection of the top 1000 features followed by a support vector machine (SVM) classifier within a leave-one-out cross-validation scheme. We did not optimize for the number features nor for the SVM classifier parameters, which we kept identical with the best setting in the original manuscript. We obtained results that are significantly above chance for discrimination, for example, between SD-fMCI and SD-aMCI. The accuracy was 84.6%, with a false-positive rate between 6.7% single-domain frontal mild cognitive impairment (SD-fMCI) and 27.3% single-domain amnesic mild cognitive impairment (SD-aMCI). As can be expected, these results are less optimistic than the reported values in the manuscript. This can be explained by a number of factors. For example, we note that the data analysis pipeline is not yet fully optimized because the time to respond to the letter by Diciotti et al was limited. We reason that this result nevertheless underlines the potential and feasibility of SVM for individual classification of MCI subtypes.

We would like to use this opportunity to further elaborate a number of current limitations of classification analysis at the individual level. Neuroimaging has been dominated by group-level statistical analyses to identify brain regions involved in certain diseases; however, such analyses do not necessarily reflect predictive value for the diagnosis of individual cases in clinical neuroradiology. Recent trends in neuroimaging data analysis are increasingly adopting tools from pattern recognition to evaluate results and develop potentially new imaging markers. This trend represents a fundamental shift in paradigm. Diciotti et al highlighted, in their letter, the issue of proper feature selection, and we would like to add a few other considerations for future development. In addition to these methodologic challenges, there are open medicolegal issues such as approval by the FDA or European Union.

1) A typical feature set extracted from MR imaging data can easily contain more than 100,000 features. In most cases, the features are related (similarity of adjacent or homologous voxels), and only a limited number will carry discriminative information. The selection of the best features is a long-standing problem in machine learning, which can be dealt with either explicitly (by a separate feature-selection step) or implicitly (by regularization in the classification method). In any case, increasing the dimensionality by adding more nondiscriminatory features increases computational demands and decreases performance. Identifying the optimal number of features (or tuning of regularization parameters) is nontrivial in practice.

2) Structural MR imaging data typically have several hundred thousand voxels, while typical single-center studies have around 20–50 individuals per group. The cross-validation technique is one frequently implemented method in such cases, with a small number of participants with respect to the size of the data (also implemented by Diciotti et al), yet it has its own limitations. Ideally, the classifier should be trained on one dataset and tested on another independent dataset to estimate the “real world” performance in clinical neuroradiology. Evidently, the available sample size for single-center studies is, in most cases, insufficient.

3) Support vector machines have been widely applied to neuroimaging data, probably because of their robustness against outliers, yet they were not specifically developed for neuroimaging data. SVM does not exploit spatial structure (ie, features can be randomly permuted without modifying the results). However, the brain has specific spatial structure so that adjacent or homologous voxels are more likely to have similar features compared with distant voxels. Introducing prior information about spatial structure in classification algorithms is an important research topic (eg, some algorithms recently proposed hierarchical clustering to regroup similar voxels and reinforce the robustness).<sup>1</sup>

4) There is substantial normal interindividual variation in brain morphometry, even in healthy volunteers (eg, up to 15% variation in cortical thickness).<sup>2</sup> Correspondingly, we could, for example, show in a previous study that there is less variation in the within-subject cortical asymmetry, and for example, discrimination between at-risk mental state and volunteers was possible only based on within-subject cortical asymmetry, yet this was impossible based on direct assessment of cortical thickness between subjects.<sup>3</sup> While this study was performed in the domain of psychosis, the principles are also applicable to neurodegenerative diseases including dementia. While most classification techniques can exploit multivariate information and, thus theoretically, can reveal discriminative information by “clever” combinations of features, the high-dimensional nature and variability of the data could benefit from incorporating domain-specific knowledge.

5) There is interindividual variation in the neurocognitive reserve, which was described already in 1968.<sup>4</sup> The same degree of clinical neurocognitive impairment can be caused by different levels of brain pathology—or from the other perspective, the same degree of brain pathology can evoke variable degrees of clinical neurocognitive impairment. This is due to individual factors such as education and social integration, which represent an important confound for classification analyses. Taking into account these factors is not obvious and should ideally be done within the classification algorithm and not as a separate preprocessing step (on the training data within the cross-validation fold).

6) MR imaging usually includes multiple pulse sequences. To increase the accuracy and, in particular, the robustness of individual classification analyses, it is probably beneficial to combine the information of multiple pulse sequences, ideally in combination with nonimaging parameters such as neuropsychologic tests, blood or CSF samples, and so forth. Determining the optimal combination of multiple domains in practice is, however, not trivial.<sup>5</sup>

7) Additional potentially confounding factors include noise in

the data, between-scanner variability, variability in data preprocessing, and patient selection, among others.

In summary, individual-level classification of neuroimaging data is an emerging field and is still hindered by fundamental limitations of the methodology, including optimal feature selection, incorporating domain knowledge into the classification, and integration of multiparametric measurements.

In addition, we believe that further methodologic developments should be based on larger datasets and multicentric studies to increase both reproducibility and predictability. Recent data-sharing initiatives such as the Alzheimer's Disease Neuroimaging Initiative,<sup>6</sup> in combination with cloud-computing power, will provide the necessary prerequisites for these developments. In the near future, we will, hopefully, see new advances to bring individual-level classification analysis to the next level to provide earlier and more accurate diagnosis and to eventually improve patient care.

## REFERENCES

1. Jenatton R, Gramfort A, Michel V, et al. **Multiscale mining of fMRI data with hierarchical structured sparsity.** *SIAM J Imaging Sci* 2012;5:835–56
2. Haug H. **Brain sizes, surfaces, and neuronal sizes of the cortex cerebri: a stereological investigation of man and his variability and a comparison with some mammals (primates, whales, marsupials, insectivores, and one elephant).** *Am J Anat* 1987;180:126–42
3. Haller S, Borgwardt SJ, Schindler C, et al. **Can cortical thickness asymmetry analysis contribute to detection of at-risk mental state and first-episode psychosis? A pilot study.** *Radiology* 2009;250:212–21
4. Tomlinson BE, Blessed G, Roth M. **Observations on the brains of non-demented old people.** *J Neurol Sci* 1968;7:331–56
5. Haller S, Lovblad KO, Giannakopoulos P. **Principles of classification analyses in mild cognitive impairment (MCI) and Alzheimer disease.** *J Alzheimers Dis* 2011;26(suppl 3):389–94
6. Mueller SG, Weiner MW, Thal LJ, et al. **Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI).** *Alzheimers Dement* 2005;1:55–66

**S. Haller**

**K.-O. Lovblad**

Department of Imaging and Medical Informatics

**P. Giannakopoulos**

Department of Mental Health and Psychiatry  
University Hospitals of Geneva and Faculty of Medicine of the  
University of Geneva  
Geneva, Switzerland

**D. Van De Ville**

Department of Radiology and Medical Informatics  
University of Geneva  
Geneva, Switzerland  
Institute of Bioengineering  
Ecole Polytechnique Fédérale de Lausanne (EPFL)  
Lausanne, Switzerland

<http://dx.doi.org/10.3174/ajnr.A3699>