Reviewer's report

Title: Refining Developmental Coordination Disorder subtyping with multivariate statistical methods

Version: 1 Date: 2 March 2012

Reviewer: Bjoern Menze

Reviewer's report:

The authors compare three different feature selection and sample clustering methods on a set of 63 samples with 49 binary observations from a clinical study on learning disabilities. They compare the permutation importance of random forest (RF), the magnitude of coefficients from a partial least squares regression (PLS), and the magnitude of the (selected) coefficients from a sparse partial least squares regression. They present subspace projections calculated from both RF (MDS on sample proximity) and PLS (subspaces spanned by scores of first few steps). They find both random forest and partial least squares to select reasonable features for their clinical task and conclude that both measures are helpful in exploring multidimensional data sets from these kinds of clinical studies.

This is an application study and one may raise the question of how the reported results will generalize beyond this specific data set. Furthermore, one should mention that RF is a nonlinear learning algorithm, while PLS seeks for a linear class separation which may lead to differences in the feature importance scores for some data sets. Eventually the authors want to comment on this. Also, except Fig. 4a, there is little detail on or discussion of the RF sample proximity that is pointed out in the abstract. Maybe the authors want to add some more detail at least on the algorithmic procedure that is used to calculate this measure.

Overall, the manuscript is easy to read, all background is well described and the authors refer to the publicly available software used. While there is clearly an issue with methodological novelty, I would see that this manuscript may serve well as introductory reference for practitioners with related clinical data.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests