

Accepted Manuscript

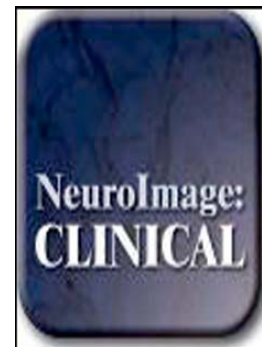
Selection bias in the reported performances of AD classification pipelines

Alex F. Mendelson, Maria A. Zuluaga, Marco Lorenzi, Brian F. Hutton,  
Sébastien Ourselin

PII: S2213-1582(16)30256-X  
DOI: doi: [10.1016/j.nicl.2016.12.018](https://doi.org/10.1016/j.nicl.2016.12.018)  
Reference: YNICKL 898

To appear in: *NeuroImage: Clinical*

Received date: 21 October 2016  
Revised date: 9 December 2016  
Accepted date: 16 December 2016



Please cite this article as: Mendelson, Alex F., Zuluaga, Maria A., Lorenzi, Marco, Hutton, Brian F., Ourselin, Sébastien, Selection bias in the reported performances of AD classification pipelines, *NeuroImage: Clinical* (2016), doi: [10.1016/j.nicl.2016.12.018](https://doi.org/10.1016/j.nicl.2016.12.018)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# Selection bias in the reported performances of AD classification pipelines

Alex F. Mendelson<sup>a,\*</sup>, Maria A. Zuluaga<sup>a</sup>, Marco Lorenzi<sup>a</sup>, Brian F. Hutton<sup>b,c</sup>, Sébastien Ourselin<sup>a,d</sup>, for the Alzheimer's Disease Neuroimaging Initiative\*\*

<sup>a</sup>*Translational Imaging Group, Centre for Medical Image Computing  
University College London, London, UK*

<sup>b</sup>*Institute of Nuclear Medicine, University College London, London, UK*

<sup>c</sup>*Centre for Medical Radiation Physics, University of Wollongong, NSW, Australia*

<sup>d</sup>*Dementia Research Centre, University College London, UK*

## Abstract

The last decade has seen a great proliferation of supervised learning pipelines for individual diagnosis and prognosis in Alzheimer's disease. As more pipelines are developed and evaluated in the search for greater performance, only those results that are relatively impressive will be selected for publication. We present an empirical study to evaluate the potential for optimistic bias in classification performance results as a result of this selection. This is achieved using a novel, resampling-based experiment design that effectively simulates the optimisation of pipeline specifications by individuals or collectives of researchers using cross validation with limited data. Our findings indicate that bias can plausibly account for an appreciable fraction (often greater than half) of the apparent performance improvement associated with the pipeline optimisation, particularly in small samples. We discuss the consistency of our findings with patterns observed in the literature and consider strategies for bias reduction and mitigation.

---

\*Corresponding author

\*\*Data used in preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.ucla.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf).

*Email address:* alexander.mendelson.11@ucl.ac.uk (Alex F. Mendelson)

Download English Version:

<https://daneshyari.com/en/article/8688693>

Download Persian Version:

<https://daneshyari.com/article/8688693>

[Daneshyari.com](https://daneshyari.com)