



# Multi-study validation of data-driven disease progression models to characterize evolution of biomarkers in Alzheimer's disease

*Damiano Archetti, Silvia Ingala, Vikram Venkatraghavan, Viktor Wottschel, Alexandra L. Young, Maura Bellio, Esther E Bron, Stefan Klein, Frederik Barkhof, Daniel C. Alexander, Neil P. Oxtoby, Giovanni B. Frisoni, Alberto Redolfi, for the Alzheimer's Disease Neuroimaging Initiative, for EuroPOND Consortium*

## Abstract:

Understanding the sequence of biological and clinical events along the course of Alzheimer's disease provides insights into dementia pathophysiology and can help participant selection in clinical trials. Our objective is to train two data-driven computational models for sequencing these events, the Event Based Model (EBM) and discriminative-EBM (DEBM), on the basis of well-characterized research data, then validate the trained models on subjects from clinical cohorts characterized by less-structured data-acquisition protocols.

Seven independent data cohorts were considered totalling 2389 cognitively normal (CN), 1424 mild cognitive impairment (MCI) and 743 Alzheimer's disease (AD) patients. The Alzheimer's Disease Neuroimaging Initiative (ADNI) data set was used as training set for the construction of disease models while a collection of multi-centric data cohorts was used as test set for validation. Cross-sectional information related to clinical, cognitive, imaging and cerebrospinal fluid (CSF) biomarkers was used.

Event sequences obtained with EBM and DEBM showed differences in the ordering of single biomarkers but according to both the first biomarkers to become abnormal were those related to CSF, followed by cognitive scores, while structural imaging showed significant volumetric decreases at later stages of the disease progression. Staging of test set subjects based on sequences obtained with both models showed good linear correlation with the Mini Mental State Examination score ( $R^2_{EBM} = 0.866$ ;  $R^2_{DEBM} = 0.906$ ). In discriminant analyses, significant differences ( $p\text{-value} \leq 0.05$ ) between the staging of subjects from training and test sets were observed in both models. No significant difference between the staging of subjects from the training and test was observed ( $p\text{-value} > 0.05$ ) when considering a subset composed by 562 subjects for which all biomarker families (cognitive, imaging and CSF) are available.

Event sequence obtained with DEBM recapitulates the heuristic models in a data-driven fashion and is clinically plausible. We demonstrated inter-cohort transferability of two disease progression models and their robustness in detecting AD phases. This is an important step towards the adoption of data-driven statistical models into clinical domain.

**NeuroImage: Clinical, 2019 Jul 23;24:101954.**

Published: 23 July 2019

<https://doi.org/10.1016/j.nicl.2019.101954>

