

Predicting Conversions from Normal to MCI and from MCI to AD with Clinical Variables in the National Alzheimer's Coordinating Center Uniform Data Set Version 3: Application of Big Data Analytics and Transition Probability Calculators

Yijiang Pang¹ | Jiayu Zhou¹ | Roger L. Albin^{2,3,4} | Mary Sano⁵ | Changyu Shen⁶ |
Walter A. Kukull^{7,8} | Hiroko H Dodge⁹

¹Michigan State University, East Lansing, MI, USA

²VA Ann Arbor Healthcare System, Ann Arbor, MI, USA

³Department of Neurology, University of Michigan, Ann Arbor, MI, USA, Ann Arbor, MI, USA

⁴Michigan Alzheimer's Disease Center, Ann Arbor, MI, USA

⁵Department of Psychiatry, Alzheimer's Disease Research Center, Icahn School of Medicine at Mount Sinai, New York, NY, USA

⁶Biogen, Inc, Cambridge, MA, USA

⁷National Alzheimer's Coordinating Center, University of Washington, Seattle, WA, USA

⁸University of Washington, Seattle, WA, USA

⁹NIA-Layton Aging & Alzheimer's Disease Center, Oregon Health & Science University, Portland, OR, USA

Correspondence

Hiroko H Dodge, NIA-Layton Aging & Alzheimer's Disease Center, Oregon Health & Science University, Portland, OR, USA.
Email: dodgeh@ohsu.edu

Abstract

Background: Clinical trials are increasingly focused on pre-manifest and early Alzheimer's disease (AD). Accurately predicting those who transit from normal to MCI or from MCI to dementia/AD (true positives) versus non-progressors (false positives) is challenging. Biomarker positivity and comorbidity information are often not sufficient to identify those who are destined to have symptom progressions. Accurate identification of progressors is important to avoid un-necessary treatment as well as to improve trial efficiency.

Method: Using only clinical variables and big data analytics, we aimed to predict those who convert from normal to MCI and from MCI to AD, estimating probabilities of these conversions with a small set of selected clinical variables. This work updates our previous work using National Alzheimer's Coordinating Center (NACC) Uniform Data Set Version 2. ⁽¹⁾ We also generated a user-friendly conversion probability calculator to estimate individual subject conversion. We used data from the NACC Uniform Data Set Version 3 and machine learning feature selection methods to identify clinical variables that predicted MCI and dementia conversion. We built a pipeline with machine learning techniques to test model performance.

Result: Using NACC data (frozen 3/2020), 577 subjects were used to predict conversions from normal to MCI within 4 years, and 538 and 882 subjects were used to predict conversion from MCI to AD within 2 years and 3 years, respectively. An exploratory analysis was also conducted for amnesic MCI (aMCI) conversion. For each prediction model, we selected 20 clinical variables optimizing predictions using Receiver Operating Characteristics Area-under-Curve (AUC). Our model achieved 82.1% prediction accuracy, 75.9% specificity, 85.1% sensitivity, and 80.5% AUC for MCI to AD conversion within 3 years. Similar results were obtained for normal to MCI conversion. The variables which ranked highest in predicting transitions include clinician's judgement of subject's meaningful cognitive impairment (conversion to MCI) and CDR community affairs item (conversion to AD in 3 years).

Conclusion: Although NACC UDS is extensive, our algorithm is based on a large sample size of potential trial participants who visited or are referred to Alzheimer's Disease Centers. The probability calculator can be used to effectively pre-screen and enrich trial participants.

Figures

Fig. 1. Prediction model performance evaluation scheme.

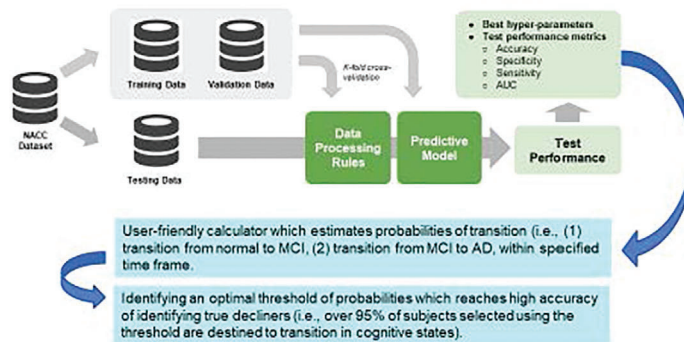


Figure 2. AUC by number of selected features.

NC2MCI: predicting conversion from normal cognition to MCI within 4 years
 NC2aMCI: predicting conversion from normal cognition to amnesic MCI within 4 years
 MCI2AD_3: predicting conversion from MCI to AD within 3 years
 MCI2AD_2: predicting conversion from MCI to AD within 2 years

