

# **Dementia risk prediction: effects of age specification and follow-up duration on model accuracy**

Md Hamidul Huque<sup>1,2,3\*</sup> and Kaarin J. Anstey<sup>1,2,3</sup>

<sup>1</sup>UNSW Ageing Futures Institute, University of New South Wales, Australia

<sup>2</sup>School of Psychology, University of New South Wales, Australia

<sup>3</sup>Brain Health and Dementia Centre, Neuroscience Research Australia, Australia

## **Abstract**

**Introduction:** We evaluated the impact of age specification and follow-up duration on the performance of three midlife risk scores: CogDrisk-ML, CAIDE, and UKBDRS.

**Methods:** Data were obtained from the ARIC, UK Biobank, and Whitehall II cohorts. Dementia risk prediction was evaluated using Cox models with three age specifications: UKBDRS (continuous), CogDrisk-ML (categorical: 40–49, 50–54, 55–59, 60–64 years), and CAIDE (categorical: <47, 47–53, >53 years). Predictive accuracy was measured using Harrell’s C-statistics.

**Results:** Continuous age produced the highest Harrell’s C-statistics (95% CI) [ARIC: 0.71 (0.68, 0.73); UK-Biobank: 0.78 (0.77, 0.79), and Whitehall II: 0.75 (0.62, 0.88)] followed by categorical age in CogDrisk-ML and CAIDE. Compared with continuous age, categorical specification reduced C-statistics, with the largest reduction for CAIDE. Shorter follow-up yielded higher but less precise C-statistics.

**Discussion:** Age specification strongly affects dementia risk model performance, with continuous age giving highest accuracy, especially with shorter follow-up.