

18 **Authorship contributions:**

- 19 • Jingjing Yang: Drafting/revision of the manuscript for content, including medical
20 writing for content; Study design; Analysis or interpretation of data
- 21 • Shahram Oveisgharan: Drafting/revision of the manuscript for content, including
22 medical writing for content; Major role in the acquisition of data; Analysis or
23 interpretation of data
- 24 • Xizhu Liu: Analysis or interpretation of data
- 25 • Robert S Wilson: Major role in the acquisition of data
- 26 • David A Bennett: Drafting/revision of the manuscript for content, including
27 medical writing for content; Major role in the acquisition of data; Study concept or
28 design; Analysis or interpretation of data
- 29 • Aron S Buchman: Drafting/revision of the manuscript for content, including
30 medical writing for content; Major role in the acquisition of data; Study concept or
31 design; Analysis or interpretation of data

32

33

34 **ABSTRACT**

35 **Background:** Alzheimer's disease is a progressive disorder without a cure. Developing
36 risk prediction models for detecting presymptomatic Alzheimer's disease using non-
37 cognitive measures is necessary to enable early interventions.

38 **Objective:** Examine if non-cognitive metrics alone can be used to construct risk models
39 to identify adults at risk for Alzheimer's dementia and cognitive impairment.

40 **Methods:** Clinical data from older adults without dementia from the Memory and Aging
41 Project (MAP, n=1179) and Religious Orders Study (ROS, n=1103) were analyzed
42 using Cox proportional hazard models to develop risk prediction models for Alzheimer's
43 dementia and cognitive impairment. Models using only non-cognitive covariates were
44 compared to models that added cognitive covariates. All models were trained in MAP,
45 tested in ROS, and evaluated by the AUC of ROC curve.

46 **Results:** Models based on non-cognitive covariates alone achieved AUC (0.800,0.785)
47 for predicting Alzheimer's dementia (3,5) years from baseline. Including additional
48 cognitive covariates improved AUC to (0.916,0.881). A model with a single covariate of
49 composite cognition score achieved AUC (0.905,0.863). Models based on non-cognitive
50 covariates alone achieved AUC (0.717,0.714) for predicting cognitive impairment (3,5)
51 years from baseline. Including additional cognitive covariates improved AUC to
52 (0.783,0.770). A model with a single covariate of composite cognition score achieved
53 AUC (0.754,0.730).

54 **Conclusion:** Risk models based on non-cognitive metrics predict both Alzheimer's
55 dementia and cognitive impairment. However, non-cognitive covariates do not provide
56 incremental predictivity for models that include cognitive metrics in predicting

57 Alzheimer's dementia, but do in models predicting cognitive impairment. Further
58 improved risk prediction models for cognitive impairment are needed.

59

60 **Keywords:** Alzheimer's disease, Cognitive aging, Mild cognitive impairment, Cohort

61 study, Cox proportional hazard model

62 INTRODUCTION

63 Alzheimer's disease is a progressive disorder that develops over years. An
64 asymptomatic stage during which Alzheimer's disease pathology (AD) accumulates is
65 followed by mild cognitive impairment (MCI), while dementia occurs later [1-3].
66 Alzheimer's disease is heterogeneous; most demented individuals transition through all
67 three clinical stages i.e., from no cognitive impairment (NCI) to MCI followed by
68 dementia. Some individual's transition directly from NCI to dementia. On the other hand,
69 some never progress beyond MCI and 1/3 with NCI prior to death also show pathologic
70 AD at death.

71 Early detection of adults with normal cognition at risk for Alzheimer's disease
72 cognitive traits is crucial for early targeted treatments [4, 5]. This underscores the
73 widespread efforts to develop risk prediction models for detecting presymptomatic
74 Alzheimer's disease. Yet, using cognitive test scores as a predictor for the incident
75 cognitive impairment (MCI or Alzheimer's dementia) may be statistically problematic as
76 the same test scores may be used as model predictors as well as for the diagnostic
77 characterization of cognitive status, which is the outcome the model seeks to identify.
78 Moreover, comprehensive cognitive testing requires in-person testing that is not widely
79 available. This highlights the potential benefits of identifying non-cognitive covariates
80 that may be obtained remotely and contain predictive information to identify cognitively
81 normal adults at risk for incident cognitive impairment and Alzheimer's dementia.

82 Alzheimer's disease is a complex disorder that not only negatively affects
83 cognition, but also adversely affects other important non-cognitive aging phenotypes [6].
84 For example, accumulation of AD is associated with body mass index (BMI) decline or

85 impaired motor function that may precede and predict incident cognitive impairment and
86 Alzheimer's dementia [7-11]. Gait velocity and variability are associated with cognitive
87 decline and gait is shown as a proxy for overall health and as an index of cognitive
88 decline [12]. These data suggest that developing a risk profile based on non-cognitive
89 covariates may facilitate the identification of at risk adults during the earliest stages of
90 Alzheimer's disease.

91 To test this hypothesis, it is crucial to examine non-cognitive and cognitive
92 covariates alone and together to compare model performance and identify the optimal
93 covariates for predicting incident Alzheimer's dementia and incident cognitive
94 impairment. This study used clinical data from older adults without dementia
95 participating in one of two prospective community-based cohort studies of aging, Rush
96 Memory and Aging Project (MAP) and the Religious Orders Study (ROS) [13]. We
97 trained models in MAP and validated model prediction performance in ROS.

98

99 **MATERIALS AND METHODS**

100 ***Participants***

101 Participants were community dwelling older persons enrolled in one of two
102 ongoing cohort studies of aging and dementia, MAP (n=1179) and ROS (n=1103).
103 Participants were enrolled without known dementia; 1742 with NCI and 540 with MCI at
104 enrollment (**Tables S1-S2**). All agreed to annual clinical evaluations and autopsy at the
105 time of death. The duration of follow up for participants ranged from 2 to 26 years, with
106 a median of 8-year and standard deviation 5.42 (**Supplementary Figure 1**). Both

107 cohorts employ a harmonized data collection battery administered by the same
108 research assistants facilitating joint analyses.

109

110 ***Cognitive Assessment and Diagnoses***

111 Detailed structured clinical examinations were administered annually to
112 participants in both studies. The neuropsychological battery includes 17 tests that
113 assess five domains of cognitive ability. Raw test scores were standardized per test
114 using baseline means and standard deviations (SDs) of both cohorts; the resulting Z-
115 scores were then averaged across the 17 cognitive tests to derive a summary
116 composite cognition score and scores for five cognitive abilities including: episodic
117 memory, semantic memory, working memory, visuospatial abilities, and perceptual
118 speed.

119 A neuropsychologist and a neurologist with expertise in dementia reviewed
120 annual cognitive testing and classified cognitive status of participants as NCI, MCI,
121 Alzheimer's dementia, or other dementias according to established NIA-AA criteria [13-
122 16]. We excluded individuals who developed non-Alzheimer's dementia from these
123 analyses.

124 This study first examined models which predicted transition from no dementia
125 (NCI or MCI) to Alzheimer's dementia (**Fig 1, Row 1**). Second, we examined models
126 which predicted transition from NCI to Alzheimer's dementia (**Fig 1, Row 2**). Third, we
127 studied the outcome of incident cognitive impairment which was defined by the first
128 annual assessment when an individual who was previously cognitively normal (NCI)

129 transitioned to cognitive impairment manifested as either MCI or Alzheimer's dementia
130 **(Fig 1, Row 3).**

131

132 ***Other Clinical Covariates***

133 Five groups of clinical variables (a total of 57) were considered potential
134 covariates in the Cox proportional hazard regression models for risk prediction (**Table 1;**
135 **Tables S3-S4**) — i) *Common Alzheimer's disease risk factors* including age, sex,
136 education, Mini Mental State Exam (MMSE) score [17] and *APOE E4* allele; ii) *Health*
137 *measures* such as blood pressure, depression, and cardiovascular diseases; iii)
138 *Medication usage* such as lipid lowering medications, and antidepressants; iv) *Variables*
139 *uniquely profiled by MAP/ROS* such as composite cognition scores based on 17
140 cognitive tests, self-reported physical activity and social network size [13]; v) *Motor and*
141 *sleep metrics* such as dexterity, hand strength, gait function, four parkinsonian signs
142 [16], and four self-reported variables about sleep quality and duration [15].

143 These 57 clinical covariates were selected by excluding variables with a high
144 proportion of missing values (>20%) and those that were highly correlated with other
145 clinical variables (correlation>0.95). Missing baseline values of longitudinal variables
146 were imputed from the corresponding values at participant's nearest past visit. A total of
147 345 individuals with missing cross-sectional variables were excluded, leaving n=1179
148 MAP and n=1103 ROS participants for these analyses. A heat map showing the
149 correlations of the covariates examined in these models is shown in **Supplementary**
150 **Figure 2.**

151

152 **Analytic Approach**

153 The annual cognitive status diagnosis and the follow up year were used to
154 identify the first occurrence of Alzheimer’s dementia or incident cognitive impairment.
155 For each participant, the year of enrollment is considered as baseline (time 0), the year
156 of first diagnosis of Alzheimer’s dementia and cognitive impairment (MCI or Alzheimer’s
157 dementia) is considered as the time when the event occurs (incident Alzheimer’s
158 dementia or incident cognitive impairment), and the last visit of participants without the
159 considered event during all follow-ups is considered as the right censored time for living
160 participants or the time of death for dead participants. Less than 10% of participants
161 developing Alzheimer’s dementia progressed directly from NCI to Alzheimer’s dementia.
162 Sample size distributions with respect to cognitive status at baseline and the cognition
163 event types are included in **Tables S1-S2**.

164 We fit separate Cox proportional hazard risk prediction models [18-21] for
165 incident Alzheimer’s dementia or incident cognitive impairment while accounting for the
166 competing risk of death. In the first stage of our analysis, we examined the prediction
167 accuracy of using non-cognitive covariates alone in the Cox models (Model A). Then,
168 we examined if adding cognitive covariates improved the model accuracy assessed.
169 Model performance, i.e., discrimination was evaluated by the prediction area under
170 curve (AUC) values of receiver operating characteristic curve (ROC) [22] for the risk of
171 developing Alzheimer’s dementia and cognitive impairment 3, 5 years from the analytic
172 baseline. AUC is equal to the probability that a classifier will rank a randomly chosen
173 individual with the outcome of interest (event or case) higher than a randomly chosen
174 individual without impaired cognition (control), assuming 'case' ranks higher than

175 'control' [22]. That is, the derived risk prediction model with higher AUC is expected to
176 predict the risk of developing Alzheimer's dementia or cognition impairment with higher
177 accuracy.

178 Next, we added the MMSE score (Model B) and then repeated the model a third
179 time adding the composite cognition score (Model C). In further analyses, we sought to
180 isolate the prediction accuracy of different combinations of cognitive metrics alone
181 including: composite cognition score (Model D), MMSE score (Model E), as well as the
182 five cognitive ability scores (Model F) used for its construction and the 17 individual
183 cognitive test scores (Model G). Since all covariates were standardized, the magnitudes
184 of covariate coefficients in the Cox models are comparable with respect to predictivity.

185 By selecting a risk score threshold corresponding the highest overall
186 classification accuracy (the proportion of correct discriminations in test ROS samples),
187 we compared the overall classification, sensitivity (the proportion of true positive
188 predictions in all test cases, i.e., true positive fraction), and specificity (the proportion of
189 true negative predictions in all test controls, i.e., 1 – false positive fraction), for Models
190 A-D. Samples with risk scores greater than the selected threshold were predicted as
191 positives or otherwise negatives. The sensitivity and specificity corresponding to a
192 selected risk score threshold reflect risk model performance at one point in the ROC
193 curves.

194 Backward variable selection was implemented during the training of Cox models
195 that consider more than one covariate. In particular, we started with the full Cox model
196 with all potential covariates; proceeded with a temporary Cox model by excluding the
197 covariate with the least significant p-value; compared their C-statistics (equivalent to

198 AUC) and conducted ANOVA test to examine if the temporary Cox model would be
199 improved over the current Cox model; accepted the temporary Cox model if the C-
200 statistic was improved in the temporary Cox model or otherwise the ANOVA test p-
201 value<0.05. We continued to exclude the covariate with the least significant p-value
202 from the accepted Cox model until the proposed temporary Cox model was not
203 accepted or when there was only one covariate remained in the model.

204 For Alzheimer's dementia, we first considered all individuals including adults with
205 either NCI or MCI at baseline, and then only individuals with NCI at baseline. Modeling
206 incident cognitive impairment considered only individuals with NCI at baseline.

207 All Cox models were trained using data from MAP participants and validated (i.e., tested)
208 with data from ROS participants [23]. We tested the respective AUC improvement in 3,
209 5 years of two models by a nonparametric test that accounts for the correlated nature of
210 the ROC curves [24]. We also evaluated the prediction accuracy, sensitivity, and
211 specificity by choosing the prediction risk score threshold with respect to the highest
212 overall accuracy.

213 ***Association Tests between Predicted Risk Scores of Alzheimer's Dementia and*** 214 ***Brain Pathologies***

215 We investigated the association between predicted risk scores (by the full Model
216 C) of NCI/MCI samples developing Alzheimer's dementia in 3, 5 years in our test ROS
217 cohort and Alzheimer's disease related dementia (ADRD) neurodegenerative and
218 cerebrovascular disease brain pathologies — Amyloid, Tangles, Global AD pathology
219 (Gpath), NIA Reagan pathology (NIA), Lewy body pathology (LewyBody), Hippocampal
220 Sclerosis (Hippo), TDP-43, Macroinfarct (Macro), Microinfarct (Micro), Atherosclerosis

221 (Arthero), Arteriolosclerosis (Arteriol), Cerebral amyloid angiopathy (CAA). Correlation
222 tests were conducted for quantitative pathologies (Amyloid, Tangles, and Gpath), and
223 two-sample t-tests were conducted for the remain dichotomized pathologies (present vs.
224 not present). Only deceased ROS participants who underwent autopsy and had profiled
225 pathologies were tested (~50%).

226

227 ***Standard Protocol Approvals, Registrations, and Patient Consents***

228 Both MAP and ROS studies were approved by an Institutional Review Board of
229 Rush University Medical Center. Written informed consent was obtained from all study
230 participants as was an Anatomical Gift Act for organ donation.

231

232 ***Data Availability***

233 All data analyzed in this study are de-identified and available to any qualified
234 investigator with application through the Rush Alzheimer's Disease Center Research
235 Resource Sharing Hub, <https://www.radc.rush.edu>, which has descriptions of the
236 studies and available data.

237

238 **RESULTS**

239 ***Risk Models Predicting Alzheimer's Dementia***

240 Overall model performance: Cox proportional hazard model trained using only non-
241 cognitive covariates showed good model performance [25] for predicting Alzheimer's
242 dementia 3, 5 years from baseline (**Fig 1, Row 1 Model A**). Stepwise addition of the

243 cognitive covariate MMSE and then composite cognition yielded further significant
244 improvement in the initial model performance of using only non-cognitive covariates
245 **(Fig 1, Row 1 Model B & Model C)**. The performance accuracy of the full model
246 including non-cognitive covariates, MMSE and composite cognition score was similar to
247 the performance of a model including composite cognition score alone (Model D) for
248 Year 3. The full model outperformed Model D for Year 5 **(Fig 1, Row 1 Model C &**
249 **Model D)**.

250

251 Important clinical predictors: As expected, composite cognition score had the largest
252 protective effect size when it was considered in the full model **(Fig 2, Row 1 Model C)**.
253 Similarly, in the model including non-cognitive and MMSE covariates **(Supplementary**
254 **Table 5)**, we observed that MMSE had the largest protective effect size, followed by
255 hand strength and *APOE* E2. In contrast, increasing age and the presence of *APOE* E4
256 had the largest effect size for increased risk of Alzheimer's dementia, when only non-
257 cognitive covariates were considered in Model A.

258

259 Cognitive metrics as predictors: To better understand the role of cognitive measures in
260 predicting Alzheimer's dementia, we compared the model performance of different
261 combinations of cognitive metrics alone for predicting Alzheimer's dementia
262 **(Supplementary Figure 3 Row 1)**. Model performance for MMSE predicting
263 Alzheimer's dementia **(Supplementary Figure 3, Row 1 Model E)** was comparable
264 with a model based only on non-cognitive measures **(Fig 1, Row 1 Model A)**. The
265 various cognitive metrics based on 17 cognitive tests alone **(Supplementary Figure 3,**

266 **Row 1 Model F and Model G**) showed similar excellent model performance that was
267 not improved by the addition of the non-cognitive covariates or MMSE. Of all the
268 cognitive indices, episodic memory had the largest protective effect size
269 **(Supplementary Figure 4, Row 1).**

270
271 Exclude baseline MCI patients: Analyses of individuals without dementia include adults
272 with NCI or MCI at baseline. In further analyses, we restricted our analyses only to
273 individuals with baseline NCI to see if the same trends were observed despite the
274 truncated heterogeneity of cognitive function by exclusion of individuals with baseline
275 MCI. Model performance for non-cognitive covariates for adults with baseline NCI alone
276 was similar to models that also included adults with baseline MCI **(Fig 1, Row 2 Model**
277 **A versus Row 1 Model A)**. There was improved model performance when cognitive
278 covariates MMSE or composite cognition scores were added to the non-cognitive
279 covariates **(Fig 1, Row 2 Model B and Model C)**. Yet, improvement in model
280 performance for predicting dementia in adults with baseline NCI alone was less marked
281 compared to models that consider adults with baseline NCI or MCI **(Fig 1, Row 2 vs.**
282 **Row 1, Model B and Model C)**.

283
284 Classification accuracy: By comparing the prediction accuracy, sensitivity, and
285 specificity corresponding to a selected risk score threshold ensures the highest overall
286 accuracy per model for Models A-D. The pattern in **Table 2** was similar to **Fig 1, Row 1**
287 **and Row 2**. That is, the full Model C performed the best with overall accuracy for Year
288 (3, 5), that outperformed Model A and Model B for Year (3, 5) when composite cognition

289 was not considered, and outperformed Model D for Year 5 when only composite
290 cognition was considered. The full Model C performed comparably with Model D for
291 Year 3.

292 ***Association between Predicted Risk Scores for Alzheimer’s Dementia and Mixed-*** 293 ***Brain Pathologies***

294 Accumulating evidence suggests that clinical Alzheimer’s dementia is most
295 commonly associated with mixed brain pathologies. We examined the association
296 between predicted risk scores of Alzheimer’s dementia by the full risk model for
297 NCI/MCI samples (**Fig 1, Row 1, Model C**) and 12 different postmortem indices of
298 ADRD pathologies in our test ROS cohort. Using the significance threshold with
299 Bonferroni correction for testing 12 pathologies ($-\log_{10}(0.05/12)$), we found that the risk
300 scores predicted by the full Model C for developing Alzheimer’s dementia in 3, 5 years
301 were associated with 5 out of 12 ADRD pathologies, including both neurodegenerative
302 pathologies (Tangles, global AD pathology, NIA Reagan) and cerebrovascular disease
303 pathologies (Atherosclerosis, Cerebral amyloid angiopathy). These results
304 (**Supplemental Figure 5**) showed that the risk scores predicted by our derived models
305 were significantly associated with mixed-brain pathologies.

306

307 ***Risk Models Predicting Cognitive Impairment***

308 Overall model performance: Cox proportional hazard model trained using only non-
309 cognitive covariates showed adequate model performance [25] for predicting cognitive
310 impairment 3 and 5 years from baseline (**Fig 1, Row 3 Model A**). Yet, discrimination for
311 predicting cognitive impairment was lower than the model predictions of Alzheimer’s

312 dementia (**Fig 1, Row 3 vs. Row 1 Model A**). Stepwise addition of MMSE and then
313 composite cognition yielded further improvement in the initial model performance (**Fig 1,**
314 **Row 3 Model B and Model C**). AUC for all models predicting cognitive impairment
315 were lower than models predicting Alzheimer's dementia (**Fig 1 and Supplementary**
316 **Figure 3**). The performance accuracy of the full model that included non-cognitive,
317 MMSE and composite cognition score showed better model performance compared to
318 models with the composite cognition alone (**Fig 1, Row 3 Model C vs. Model D**).

319
320 Important clinical predictors: By examining the effect sizes in the Cox model using non-
321 cognitive and MMSE covariates (**Fig2; Supplementary Table 5**), we observed that
322 MMSE had the largest protective effect size and age had the largest effect size for
323 increased risk of cognitive impairment. Composite cognition still had the largest
324 protective effect for reducing the risk of cognitive impairment when it was considered in
325 the full model. We observed that age and *APOE E4* were associated with an increased
326 risk of both Alzheimer's dementia and cognitive impairment. Better motor function was
327 associated with reduced risk of Alzheimer' dementia (hand strength and gait) and
328 cognitive impairment (hand strength and parkinsonism) (**Fig 2; Supplementary Table**
329 **5**).

330
331 Cognitive function metrics: Next, we compared model performance for several cognitive
332 metrics alone for predicting cognitive impairment. Model performance for MMSE alone
333 (**Supplementary Figure 3, Row3 Model E**) was poor and significantly lower than a
334 model based on all the non-cognitive covariates together (**Fig 1, Row 3 Model A**).

335 Models using scores of five cognitive abilities (**Supplementary Figure 3, Row3 Model**
336 **F**) or using the 17 individual cognitive test scores (**Supplementary Figure 3, Row3**
337 **Model G**) had comparable prediction AUC as the model using composite cognition
338 alone (**Supplementary Figure 3, Row3 Model D**). These results together with the joint
339 modeling of cognitive and non-cognitive covariates (**Fig1, Row 3**) show that non-
340 cognitive clinical covariates contain additional predictive information for predicting
341 cognitive impairment not be accounted for by using cognitive covariates alone.

342
343 Classification accuracy: By comparing the prediction accuracy, sensitivity, and
344 specificity corresponding to a selected risk score threshold ensures the highest overall
345 accuracy per model for Models A-D (**Table 2**). These models show a similar pattern to
346 **Fig 1, Row 3**. That is, the full Model C performed the best with the highest overall
347 accuracy and specificity for Year (3, 5) performing better than models that did not
348 include global cognition score (A and B) or a model with global cognition score alone
349 (Model D). So, joint models of non-cognitive and cognitive measures provide better
350 predictions of cognitive impairment than cognitive measures alone.

351

352 **Discussion**

353 ***Risk models for incident Alzheimer's dementia***

354 This study used a wide range of non-cognitive covariates to construct
355 multifactorial risk models that as hypothesized yielded good model performance for
356 predicting incident Alzheimer's dementia. An important aspect of this study was that we

357 examined model performance achieved by non-cognitive covariates and cognitive
358 function metrics alone and jointly in adults with no dementia. This analytic approach
359 allows us to isolate the potential improvement in model performance due to non-
360 cognitive or cognitive covariates. Our results highlight the effectiveness of cognitive
361 measures alone in predicting Alzheimer's dementia.

362 While risk models like the Framingham Risk Score have been developed to
363 predict vascular events [26], there is no widely accepted risk model for the prediction of
364 Alzheimer's dementia or cognitive impairment [5, 27-32] (**Supplementary Table 6**). In
365 addition, the aging field has not converged on a parsimonious set of dementia risk
366 factors that must be included when developing risk models for cognitive outcomes. Thus,
367 many studies have included cognitive metrics together with non-cognitive covariates [33,
368 34] and many other studies have not included cognitive measures [5, 27, 28, 30, 35].
369 Given the varied covariates employed in prior publications, it is not surprising that model
370 discrimination has ranged from an AUC of 0.55 to 0.95 for predicting Alzheimer's
371 dementia. However, even studies reporting models with adequate discrimination have
372 not usually replicated these findings in independent cohorts. Moreover, there has not
373 been a systematic effort to clarify the independent roles of non-cognitive versus
374 cognitive metrics for predicting Alzheimer's dementia as well as consensus about
375 whether to include cognitive measures as both predictors and outcome in the same
376 models. Nonetheless, review of these past studies suggests that overall, model
377 performance for predicting Alzheimer's dementia in very old adults was lower when risk
378 models were based on non-cognitive metrics alone [28].

379 Our results, validated in an independent cohort, suggest that risk models using
380 non-cognitive covariates show good model performance and predict Alzheimer's
381 dementia. We also showed that this good performance was further improved by adding
382 cognitive scores. Yet, further analyses showed that the risk models based on composite
383 cognitive metrics alone achieved better model performance than the model based on
384 non-cognitive covariates and MMSE for Alzheimer's dementia (**Fig 1, Row 1**). By
385 restricting our modeling only to adults with NCI, we showed that non-cognitive variables
386 add additional predictive information that cannot be accounted for by the composite
387 cognitive scores alone for predicting Alzheimer's dementia (**Fig 1, Row 2**). These
388 results may have important clinical consequences for public health efforts to identify
389 adults with presymptomatic Alzheimer's disease, i.e., NCI at risk for future dementia.
390 Further work is needed to identify non-cognitive covariates or AD biomarkers that may
391 improve risk models in adults with NCI at the earliest stages of Alzheimer's disease.

392

393 ***Risk models for incident cognitive impairment***

394 Transition from NCI to MCI, an earlier stage in the trajectory of Alzheimer's
395 disease, occurred in almost all participants (>90%) who developed Alzheimer's
396 dementia in MAP/ROS studies. Since the biology underlying earlier and later stages of
397 Alzheimer's disease may vary, their respective risk models may be different [3].
398 Compared to studying Alzheimer's dementia, larger sample size and wider range of
399 clinical variables are needed for studying cognitive impairment that includes transition
400 from NCI to either MCI or Alzheimer's dementia. Thus, there are far fewer studies which
401 have developed risk models for predicting cognitive impairment as compared to

402 dementia and even fewer have validated findings in independent samples. Model
403 performances reported for predicting cognitive impairment have varied widely
404 (**Supplementary Table 7**) [34, 36-38]. We are unaware of studies that isolated the
405 independent and joint contributions of non-cognitive and cognitive metrics for model
406 discrimination of incident cognitive impairment.

407 Our study found that risk models for cognitive impairment based on only non-
408 cognitive covariates showed adequate model performance but consistently poorer than
409 prediction models for Alzheimer's dementia (**Fig 1 Row 2 vs Row 3**). While adding
410 cognitive covariates in addition to non-cognitive covariates improved model
411 performance predicting cognitive impairment, overall performance never equaled model
412 performance for predicting Alzheimer's dementia. These findings highlight the
413 importance of further studies to identify additional clinical metrics or biomarkers to
414 improve the prediction of cognitive impairment.

415

416 ***Risk models for cognitive impairment versus Alzheimer's dementia***

417 An additional strength of the current study design is that we compared the
418 performances of models predicting cognitive impairment and Alzheimer's dementia
419 using the same covariates and individuals with NCI at baseline. In addition, we validated
420 our results in an independent sample of older adults. Model performance for predicting
421 Alzheimer's dementia was much better than models predicting cognitive impairment.
422 These findings may derive from the differences in the symptoms and diagnoses of MCI
423 versus Alzheimer's dementia. That is, discriminating symptoms of Alzheimer's dementia
424 from NCI/MCI is generally much easier than discriminating symptoms of MCI from NCI.

425 There is more random error and instability in the diagnosis of MCI compared to
426 dementia. Many prior studies have reported lower sensitivity and specificity for cognitive
427 instruments used to identify MCI compared to Alzheimer's dementia [39]. Individuals
428 with MCI show more fluctuations between NCI and MCI during repeated cognitive
429 testing. In contrast, once individuals receive a diagnosis of dementia, they are less likely
430 to revert back to NCI/MCI on repeated testing.

431 Yet, varied model performance for predicting cognitive impairment and
432 Alzheimer's dementia may inform differences in the respective biology underlying these
433 Alzheimer's clinical traits [40]. Additionally, the associations between a specific risk
434 factor with dementia may vary over time. For example, high BMI in mid-life may be
435 related to an increased risk of dementia, while high BMI in much older adults may be
436 related to decreased risk of dementia [41]. Analogously, in earlier stages of Alzheimer's
437 disease, the burden and location of AD may lead to impaired non-cognitive function like
438 motor function to a greater extent than cognition. In contrast, during its later stages the
439 widespread accumulation of AD throughout cognitive brain regions may have a larger
440 pervasive negative impact on cognition. While cognitive and motor decline in older adult
441 are both commonly associated with mixed-brain pathologies, different combinations of
442 pathologies may be associated with cognitive and motor phenotypes in the same
443 individual [6, 11]. This might account for better performance achieved by cognitive
444 metrics for Alzheimer's dementia and the lack of incremental validity of non-cognitive
445 covariates when considered together with cognitive metrics for Alzheimer's dementia.
446 This might also account for the incremental predictivity of non-cognitive covariates in
447 addition to cognitive metrics for incident cognitive impairment.

448

449 ***Implications of the current findings***

450 The reconceptualization of Alzheimer’s disease as a chronic progressive disorder
451 that develops over many years suggests that treatments during its earliest stage when
452 adults manifest normal cognitive function may offer the best chance of preventing future
453 cognitive impairment. However, this opportunity requires the identification of at-risk
454 older adults who are likely to benefit from early interventions. Our current findings lend
455 further support for accumulating evidence that Alzheimer’s disease affects diverse
456 cognitive and other non-cognitive aging phenotypes such as physical frailty and
457 sarcopenia [6, 10, 12, 42]. Thus, our findings are timely as they inform on how to
458 optimize the construction of models to identify at risk older adults by including a wider
459 array of motor, cognitive and behavioral metrics [1, 2, 43, 44]. To date there is no
460 consensus about whether cognitive function metrics can or should be included together
461 with non-cognitive function metrics for predicting Alzheimer’s dementia. The current
462 results demonstrate the importance of addressing this issue since including cognitive
463 function metrics can impact model performance and the value added for non-cognitive
464 measures used to construct risk models for different Alzheimer’s cognitive phenotypes.

465 Examining important predictive covariates selected in our Cox risk models
466 highlights the importance of motor function (e.g., hand strength and Parkinsonian gait
467 score) as one of the most important non-cognitive covariates in predicting both incident
468 Alzheimer’s dementia and incident cognitive impairment. Although our previous study
469 showed that when cognitive metrics are considered motor function did not show
470 incremental predictivity for Alzheimer’s dementia [45], this study showed that motor

471 metrics provide additional predictive information for both Alzheimer’s dementia and
472 cognitive impairment (**Fig 2; Supplementary Table 5**). Motor planning and ongoing
473 attention is necessary for all movement. Thus, the fact that “we must think first and then
474 move next” may explain why motor metrics predict cognitive impairment. Yet
475 conventional motor testing only quantifies movement and does not assess specific
476 cognitive indices underlying motor planning or executive and attentional resources
477 crucial for movement execution. Instrumented gait testing with unobtrusive sensors can
478 be used to quantify cognitive mobility metrics which precede movement [46, 47]. Further
479 work is needed to determine if adding additional cognitive mobility metrics may improve
480 predictions of adults at risk for Alzheimer’s dementia and cognitive impairment.

481 Also, broadening the motor metrics used to predict cognitive impairment might
482 help circumvent statistical concerns about confounding predictors with model outcomes
483 as cognitive mobility metrics are not part of current cognitive batteries used to document
484 dementia or cognitive impairment. While there are many knowledge gaps in our
485 understanding of the trajectory for the accumulation of AD pathology, the time course
486 for β -amyloid and tau-tangles during the early and later stages of Alzheimer’s disease
487 may vary; so fluid biomarkers of AD traits might be particularly helpful for improving
488 model performance [34, 48]. Finally, leveraging the rapid advances in machine learning
489 suggests that it may be possible to impute and include underlying AD pathologies to
490 improve model performances [23, 49-53].

491

492 ***Strength & Limitations***

493 The study has several strengths that lend confidence for the current findings. All
494 subjects were recruited from the community, underwent an annual detailed clinical
495 evaluation, and were determined to be free of dementia at study entry. Large numbers
496 of men and women underwent annual assessments, and follow-up rates were very high.
497 Uniform, structured procedures were followed that enhanced stability of diagnoses
498 across time, space, and examiners. Diverse non-cognitive covariates were examined.
499 An important strength of the current study design is that we compared the performances
500 of models predicting cognitive impairment and Alzheimer's dementia in the same
501 individuals with identical covariates [54]. Moreover, we were able to replicate our results
502 in an independent sample of older adults who underwent the same assessments and
503 testing by the same staff. However, further studies will be needed due to limitations.
504 Participants were predominantly Americans of European descent and have higher than
505 average levels of education, so our findings will need to be replicated in more diverse
506 populations. Brain imaging and fluid biomarkers were not examined. While we
507 employed the NIA-AA criterion for clinical diagnosis of Alzheimer's dementia, we
508 recognize that this diagnosis is often associated with mixed brain pathologies as shown
509 in our analysis of decedents who underwent brain autopsy. The particular cognitive
510 battery employed is more extensive than what might be available outside of a research
511 setting highlighting the importance of our findings using the MMSE.

512

513

514 **Acknowledgments**

515 This work was supported by the National Institute of Health R35GM138313,
516 P30AG10161, K01AG054700, R01AG15819, R01AG17917, R01AG56352; the Illinois
517 Department of Public Health; and the Robert C. Borwell Endowment Fund. The funding
518 organizations had no role in the design or conduct of the study; collection, management,
519 analysis, or interpretation of the data; or preparation, review, or approval of the
520 manuscript. We are deeply indebted to all participants who contributed their data and
521 agreed to autopsy at the time of their death. We are thankful to the staff at the Rush
522 Alzheimer's Disease Center.

523

524 **Disclosure Statement**

525 The authors have no conflict of interest to report.

526 References

- 527 [1] Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E (2011) Alzheimer's
528 disease. *Lancet* **377**, 1019-1031.
- 529 [2] Burns A, Iliffe S (2009) Alzheimer's disease. *BMJ* **338**, b158.
- 530 [3] Jack CR, Jr., Albert M, Knopman DS, McKhann GM, Sperling RA, Carrillo MC, Thies B,
531 Phelps CH (2011) Introduction to revised criteria for the diagnosis of Alzheimer's disease:
532 National Institute on Aging and the Alzheimer's Association workgroup. *Alzheimers*
533 *Dement* **7**, 256-262.
- 534 [4] Barnes DE, Lee SJ (2011) Predicting Alzheimer's risk: why and how? *Alzheimer's Research*
535 *& Therapy* **3**, 33.
- 536 [5] Reitz C, Tang MX, Schupf N, Manly JJ, Mayeux R, Luchsinger JA (2010) A summary risk
537 score for the prediction of Alzheimer disease in elderly persons. *Arch Neurol* **67**, 835-841.
- 538 [6] Albers MW, Gilmore GC, Kaye J, Murphy C, Wingfield A, Bennett DA, Boxer AL, Buchman
539 AS, Cruickshanks KJ, Devanand DP, Duffy CJ, Gall CM, Gates GA, Granholm AC, Hensch T,
540 Holtzer R, Hyman BT, Lin FR, McKee AC, Morris JC, Petersen RC, Silbert LC, Struble RG,
541 Trojanowski JQ, Verghese J, Wilson DA, Xu S, Zhang LI (2015) At the interface of sensory
542 and motor dysfunctions and Alzheimer's disease. *Alzheimers Dement* **11**, 70-98.
- 543 [7] Buchman AS, Schneider JA, Wilson RS, Bienias JL, Bennett DA (2006) Body mass index in
544 older persons is associated with Alzheimer disease pathology. *Neurology* **67**, 1949-1954.
- 545 [8] Buchman AS, Wilson RS, Bienias JL, Shah RC, Evans DA, Bennett DA (2005) Change in
546 body mass index and risk of incident Alzheimer disease. *Neurology* **65**, 892-897.
- 547 [9] Yu L, Boyle PA, Leurgans SE, Wilson RS, Bennett DA, Buchman AS (2019) Incident
548 Mobility Disability, Mild Cognitive Impairment, and Mortality in Community-Dwelling
549 Older Adults. *Neuroepidemiology*, 1-8.
- 550 [10] Beeri MS, Leurgans SE, Delbono O, Bennett DA, Buchman AS (2021) Sarcopenia is
551 associated with incident Alzheimer's dementia, mild cognitive impairment, and cognitive
552 decline. *J Am Geriatr Soc*.
- 553 [11] Buchman AS, Bennett DA (2022) Mixed Neuropathologies, Neural Motor Resilience and
554 Target Discovery for Therapies of Late-Life Motor Impairment. *Front. Hum. Neurosci.* **16**,
555 853330.
- 556 [12] MacDonald SW, Hundza S, Love JA, DeCarlo CA, Halliday DW, Brewster PW, Lukyn TV,
557 Camicioli R, Dixon RA (2017) Concurrent Indicators of Gait Velocity and Variability Are
558 Associated with 25-Year Cognitive Change: A Retrospective Longitudinal Investigation.
559 *Front Aging Neurosci* **9**, 17.
- 560 [13] Bennett DA, Buchman AS, Boyle PA, Barnes LL, Wilson RS, Schneider JA (2018) Religious
561 Orders Study and Rush Memory and Aging Project. *J Alzheimers Dis* **64**, S161-S189.
- 562 [14] Bennett DA, Wilson RS, Schneider JA, Evans DA, Beckett LA, Aggarwal NT, Barnes LL, Fox
563 JH, Bach J (2002) Natural history of mild cognitive impairment in older persons.
564 *Neurology* **59**, 198-205.
- 565 [15] Park M, Buchman AS, Lim AS, Leurgans SE, Bennett DA (2014) Sleep complaints and
566 incident disability in a community-based cohort study of older persons. *Am J Geriatr*
567 *Psychiatry* **22**, 718-726.

- 568 [16] Buchman AS, Yu L, Oveisgharan S, Petyuk VA, Tasaki S, Gaiteri C, Wilson RS, Grodstein F,
569 Schneider JA, Klein H-U, De Jager PL, Bennett DA (2021) Cortical proteins may provide
570 motor resilience in older adults. *Scientific Reports* **11**, 11311.
- 571 [17] Tombaugh TN, McIntyre NJ (1992) The mini-mental state examination: a comprehensive
572 review. *J Am Geriatr Soc* **40**, 922-935.
- 573 [18] Cox DR (1972) Regression models and life - tables. *Journal of the Royal Statistical*
574 *Society: Series B (Methodological)* **34**, 187-202.
- 575 [19] Benichou J, Gail MH (1990) Estimates of absolute cause-specific risk in cohort studies.
576 *Biometrics* **46**, 813-826.
- 577 [20] Gerds TA, Scheike TH, Andersen PK (2012) Absolute risk regression for competing risks:
578 interpretation, link functions, and prediction. *Stat Med* **31**, 3921-3930.
- 579 [21] Ozenne B, Sørensen AL, Scheike T, Torp-Pedersen C, Gerds TA (2017) riskRegression:
580 predicting the risk of an event using Cox regression models. *The R Journal* **9**, 440-460.
- 581 [22] Linden A (2006) Measuring diagnostic and predictive accuracy in disease management:
582 an introduction to receiver operating characteristic (ROC) analysis. *J Eval Clin Pract* **12**,
583 132-139.
- 584 [23] Hastie T, Tibshirani R, Friedman JS (2009) *The Elements of Statistical Learning*, Springer-
585 Verlag New York.
- 586 [24] DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or
587 more correlated receiver operating characteristic curves: a nonparametric approach.
588 *Biometrics* **44**, 837-845.
- 589 [25] Mandrekar JN (2010) Receiver operating characteristic curve in diagnostic test
590 assessment. *J Thorac Oncol* **5**, 1315-1316.
- 591 [26] D'Agostino RB, Sr., Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, Kannel WB
592 (2008) General cardiovascular risk profile for use in primary care: the Framingham Heart
593 Study. *Circulation* **117**, 743-753.
- 594 [27] Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J (2006) Risk
595 score for the prediction of dementia risk in 20 years among middle aged people: a
596 longitudinal, population-based study. *Lancet Neurol* **5**, 735-741.
- 597 [28] Walters K, Haroon S, Petersen I, Iliffe S, Omar RZ, Nazareth I, Rait G (2016) Predicting
598 dementia risk in primary care: development and validation of the Dementia Risk Score
599 using routinely collected data. *BMC Med* **14**, 6.
- 600 [29] Vos SJB, van Boxtel MPJ, Schiepers OJG, Deckers K, de Vugt M, Carrière I, Dartigues JF,
601 Peres K, Artero S, Ritchie K, Galluzzo L, Scafato E, Frisoni GB, Huisman M, Comijs HC,
602 Sacuiu SF, Skoog I, Irving K, O'Donnell CA, Verhey FRJ, Visser PJ, Köhler S (2017)
603 Modifiable Risk Factors for Prevention of Dementia in Midlife, Late Life and the Oldest-
604 Old: Validation of the LIBRA Index. *J Alzheimers Dis* **58**, 537-547.
- 605 [30] Li J, Ogronik M, Devine S, Auerbach S, Wolf PA, Au R (2018) Practical risk score for 5-,
606 10-, and 20-year prediction of dementia in elderly persons: Framingham Heart Study.
607 *Alzheimers Dement* **14**, 35-42.
- 608 [31] Anstey KJ, Cherbuin N, Herath PM, Qiu C, Kuller LH, Lopez OL, Wilson RS, Fratiglioni L
609 (2014) A self-report risk index to predict occurrence of dementia in three independent
610 cohorts of older adults: the ANU-ADRI. *PLoS One* **9**, e86141.

- 611 [32] van Middelaar T, Hoevenaar-Blom MP, van Gool WA, Moll van Charante EP, van Dalen
612 JW, Deckers K, Kohler S, Richard E (2018) Modifiable dementia risk score to study
613 heterogeneity in treatment effect of a dementia prevention trial: a post hoc analysis in
614 the preDIVA trial using the LIBRA index. *Alzheimers Res Ther* **10**, 62.
- 615 [33] Belleville S, Fouquet C, Hudon C, Zomahoun HTV, Croteau J, Consortium for the Early
616 Identification of Alzheimer's d-Q (2017) Neuropsychological Measures that Predict
617 Progression from Mild Cognitive Impairment to Alzheimer's type dementia in Older
618 Adults: a Systematic Review and Meta-Analysis. *Neuropsychol Rev* **27**, 328-353.
- 619 [34] Albert M, Zhu Y, Moghekar A, Mori S, Miller MI, Soldan A, Pettigrew C, Selnes O, Li S,
620 Wang MC (2018) Predicting progression from normal cognition to mild cognitive
621 impairment for individuals at 5 years. *Brain* **141**, 877-887.
- 622 [35] Lin HR, Tsuji T, Kondo K, Imanaka Y (2018) Development of a risk score for the prediction
623 of incident dementia in older adults using a frailty index and health checkup data: The
624 JAGES longitudinal study. *Prev Med* **112**, 88-96.
- 625 [36] Pankratz VS, Roberts RO, Mielke MM, Knopman DS, Jack CR, Jr., Geda YE, Rocca WA,
626 Petersen RC (2015) Predicting the risk of mild cognitive impairment in the Mayo Clinic
627 Study of Aging. *Neurology* **84**, 1433-1442.
- 628 [37] Steenland K, Zhao L, John SE, Goldstein FC, Levey A, Alvaro A, Alzheimer's Disease
629 Neuroimaging I (2018) A 'Framingham-like' Algorithm for Predicting 4-Year Risk of
630 Progression to Amnesic Mild Cognitive Impairment or Alzheimer's Disease Using
631 Multidomain Information. *J Alzheimers Dis* **63**, 1383-1393.
- 632 [38] Fouladvand S, Mielke MM, Vassilaki M, Sauver JS, Petersen RC, Sohn S (2019) Deep
633 Learning Prediction of Mild Cognitive Impairment using Electronic Health Records.
634 *Proceedings (IEEE Int Conf Bioinformatics Biomed)* **2019**, 799-806.
- 635 [39] Patnode CD, Perdue LA, Rossom RC, Rushkin MC, Redmond N, Thomas RG, Lin JS (2020)
636 Screening for Cognitive Impairment in Older Adults: Updated Evidence Report and
637 Systematic Review for the US Preventive Services Task Force. *Jama* **323**, 764-785.
- 638 [40] Zammit AR, Yu L, Petyuk V, Schneider JA, De Jaeger PL, Klein H-U, Bennett DA, Buchman
639 AS (2022) Aggregation of cortical resilience proteins yields person-specific indices of
640 cognitive resilience. *Neurology* **In Press**.
- 641 [41] García-Ptacek S, Faxén-Irving G, Cermáková P, Eriksdotter M, Religa D (2014) Body mass
642 index in dementia. *Eur J Clin Nutr* **68**, 1204-1209.
- 643 [42] Buchman AS, Boyle PA, Wilson RS, Tang Y, Bennett DA (2007) Frailty is Associated With
644 Incident Alzheimer's Disease and Cognitive Decline in the Elderly. *Psychosomatic
645 Medicine* **69**, 483-489.
- 646 [43] Jack CR, Bennett DA, Blennow K (2018) 2018 NIA-AA research framework to investigate
647 the Alzheimer's disease continuum.
- 648 [44] Montero-Odasso M, Pieruccini-Faria F, Ismail Z, Li K, Lim A, Phillips N, Kamkar N, Sarquis-
649 Adamson Y, Speechley M, Theou O, Verghese J, Wallace L, Camicioli R (2020) CCCDTD5
650 recommendations on early non cognitive markers of dementia: A Canadian consensus.
651 *Alzheimers Dement (N Y)* **6**, e12068.
- 652 [45] Beerli MS, Leurgans SE, Bennett DA, Barnes LL, Buchman AS (2021) Diverse Motor
653 Performances Are Related to Incident Cognitive Impairment in Community-Dwelling
654 Older Adults. *Front Aging Neurosci* **13**, 717139.

- 655 [46] Gazit E, Buchman AS, Dawe R, Curran TA, Mirelman A, Giladi N, Hausdorff JM (2020)
656 What happens before the first step? A New Approach to Quantifying Gait Initiation
657 Using a Wearable Sensor. *Gait Posture* **76**, 128-135.
- 658 [47] Bayot M, Dujardin K, Tard C, Defebvre L, Bonnet CT, Allart E, Delval A (2018) The
659 interaction between cognition and motor control: A theoretical framework for dual-task
660 interference effects on posture, gait initiation, gait and turning. *Neurophysiol Clin* **48**,
661 361-375.
- 662 [48] Rosenberg A, Solomon A, Jelic V, Hagman G, Bogdanovic N, Kivipelto M (2019)
663 Progression to dementia in memory clinic patients with mild cognitive impairment and
664 normal beta-amyloid. *Alzheimers Res Ther* **11**, 99.
- 665 [49] Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. *Journal of*
666 *the Royal Statistical Society: Series B (Statistical Methodology)* **67**, 301-320.
- 667 [50] Nagpal S, Meng X, Epstein MP, Tsoi LC, Patrick M, Gibson G, De Jager PL, Bennett DA,
668 Wingo AP, Wingo TS, Yang J (2019) TIGAR: An Improved Bayesian Tool for
669 Transcriptomic Data Imputation Enhances Gene Mapping of Complex Traits. *Am J Hum*
670 *Genet* **105**, 258-266.
- 671 [51] Luningham JM, Chen J, Tang S, De Jager PL, Bennett DA, Buchman AS, Yang J (2020)
672 Bayesian Genome-wide TWAS Method to Leverage both cis- and trans-eQTL Information
673 through Summary Statistics. *Am J Hum Genet* **107**, 714-726.
- 674 [52] Tang S, Buchman AS, De Jager PL, Bennett DA, Epstein MP, Yang J (2021) Novel
675 Variance-Component TWAS method for studying complex human diseases with
676 applications to Alzheimer's dementia. *PLoS Genet* **17**, e1009482.
- 677 [53] Wingo AP, Liu Y, Gerasimov ES, Gockley J, Logsdon BA, Duong DM, Dammer EB, Robins C,
678 Beach TG, Reiman EM, Epstein MP, De Jager PL, Lah JJ, Bennett DA, Seyfried NT, Levey
679 AI, Wingo TS (2021) Integrating human brain proteomes with genome-wide association
680 data implicates new proteins in Alzheimer's disease pathogenesis. *Nat Genet* **53**, 143-
681 146.
- 682 [54] Licher S, Yilmaz P, Leening MJG, Wolters FJ, Vernooij MW, Stephan BCM, Ikram MK,
683 Ikram MA (2018) External validation of four dementia prediction models for use in the
684 general community-dwelling population: a comparative analysis from the Rotterdam
685 Study. *Eur J Epidemiol* **33**, 645-655.

687

688 **Table 1. Select clinical characteristics of the analytic cohort at baseline (n=1179 in**
 689 **MAP; n=1103 in ROS).**

Group	Variable	Mean (SD) or N (%)
Demographics and Common Alzheimer’s Risk Factors	Mild cognitive impairment (N)	541 (23.7)
	Follow-up years (2-26)	8 (5.42)
	Age (years, 53-102)	77.4 (7.51)
	Sex (males, N)	577 (25.2)
	Education (years, 0-30)	16.5 (3.65)
	BMI (9.09, 62.91)	27.52 (5.50)
	MMSE Score (18-30)	28.37 (1.73)
	APOE E2 (carriers, N)	330 (14.4)
	APOE E4 (carriers, N)	513 (22.4)
Covariates Uniquely Profiled by ROS/MAP	Composite cognition test Z-score (-1.95, 1.47)	0.32 (0.51)
	Physical activity (0-35)	3.19 (3.79)
	Social network size (0, 162)	7.58 (7.52)
	Maternal education (years, 0-20)	9.66 (3.79)
	Paternal education (years, 0-24)	9.60 (4.26)
	Anxiety (0-10)	1.28 (1.69)
	Proneness to psychological distress (0-42)	15.69 (6.44)
	Tendency to be social, active, and optimistic (5, 24)	15.52 (3.14)
	Early life socioeconomic status Z-score (-2.72, 2.15)	0.02 (0.74)
	Anger (0-10)	1.76 (1.88)
	TOMM40_S (N)	1,126 (49.2)
	TOMM40_L (N)	371 (16.2)
TOMM40_VL (N)	1,114 (48.7)	
Motor Metrics	Dexterity (0, 1.85)	1.02 (0.16)
	Gait function (0, 2.13)	1.05 (0.27)
	Hand strength (0, 2.36)	1.07 (0.30)
	Bradykinesia (0, 85)	8.82 (10.6)
	Parkinsonian gait (0, 85.18)	12.16 (13.45)
	Rigidity (0, 60)	2.15 (5.88)
	Tremor (0, 69.69)	2.23 (5.02)
Self-reported Sleep Metrics	Sleep latency (1, 9)	3.63 (1.14)
	Sleep consolidation (1, 8)	2.92 (1.31)
	Daytime sleepiness (1, 5)	3.21 (1.33)
	Sleep quality (1, 5)	2.10 (1.05)

690
 691
 692

693 **Table 2. Prediction accuracy (with 95% confidence interval), sensitivity, and**
 694 **specificity with respect to selected risk score thresholds that ensure the highest**
 695 **overall accuracy.** Values in this table are reflecting risk model prediction performance
 696 at one point in the ROC curves as shown in Fig1, with corresponding risk score
 697 thresholds. Samples with predicted risk scores greater than the selected threshold were
 698 considered as Predicted Positives (incident Alzheimer’s dementia or incident cognitive
 699 impairment), otherwise Predicted Negatives (not developing Alzheimer’s dementia or
 700 NCI).

		Model A. Non-Cognitive Covariates		Model B. MMSE + Non-Cognitive Covariates		Model C. MMSE + Composite Cognition + Non- Cognitive Covariates		Model D. Composite Cognition	
		Y3	Y5	Y3	Y5	Y3	Y5	Y3	Y5
NCI/MCI -> Alzheimer’s Dementia	Accuracy^a (95% CI)	0.664 (0.637, 0.694)	0.616 (0.586, 0.644)	0.761 (0.735, 0.786)	0.707 (0.679, 0.734)	0.856 (0.834, 0.876)	0.816 (0.792, 0.838)	0.865 (0.844, 0.885)	0.784 (0.758, 0.808)
	Sensitivity^b	0.810	0.804	0.810	0.804	0.810	0.804	0.810	0.804
	Specificity^c	0.655	0.591	0.757	0.694	0.860	0.817	0.870	0.781
NCI -> Alzheimer’s Dementia	Accuracy^a (95% CI)	0.688 (0.655, 0.719)	0.701 (0.669, 0.732)	0.790 (0.761, 0.818)	0.695 (0.662, 0.726)	0.818 (0.790, 0.844)	0.763 (0.733, 0.791)	0.806 (0.778, 0.832)	0.645 (0.611, 0.677)
	Sensitivity^b	0.800	0.804	0.800	0.804	0.800	0.804	0.800	0.804
	Specificity^c	0.685	0.695	0.790	0.689	0.818	0.761	0.806	0.636
NCI -> Cognitive Impairment	Accuracy^a (95% CI)	0.650 (0.617, 0.682)	0.635 (0.601, 0.667)	0.662 (0.629, 0.693)	0.648 (0.614, 0.680)	0.701 (0.669, 0.732)	0.685 (0.679, 0.716)	0.672 (0.639, 0.703)	0.657 (0.624, 0.689)
	Sensitivity^b	0.700	0.700	0.700	0.700	0.700	0.700	0.700	0.700
	Specificity^c	0.635	0.604	0.650	0.623	0.702	0.678	0.664	0.636

- 701 a. Accuracy= (# True Positive Predictions + # True Negative Predictions) / (# of test samples)
 702 b. Sensitivity = (# True Positive Predictions) / (# Positives in test samples) = True Positive Fraction
 703 c. Specificity = (# True Negative Predictions) / (# Negatives in test samples) = 1 - False Positive Fraction
 704
 705

706 **Figure Legends**

707

708 **Fig 1. ROC plots of risk prediction for incident Alzheimer’s dementia and incident**
709 **cognitive impairment with non-cognitive and cognitive covariates alone and together.**

710 Model A uses 55 non-cognitive covariates; Model B 55 non-cognitive covariates + MMSE;

711 Model C 55 non-cognitive covariates+ MMSE+ Composite Cognition; Model D uses only

712 Composite Cognition.

713 **ROW 1 No dementia (NCI or MCI) at baseline and incident Alzheimer’s Dementia:** Model B

714 vs A showed significant improvement with p-values= $(4.85 \times 10^{-4}, 1.09 \times 10^{-4})$ for Year (3, 5).

715 Model C vs B showed significant improvement with p-values= $(7.55 \times 10^{-7}, 9.15 \times 10^{-6})$ for Year

716 (3, 5). Model C vs D for Year 3 (p-value = 0.156) but significant improvement for Year 5 (p-value

717 = 0.033).

718 **ROW 2 NCI at baseline and incident Alzheimer’s Dementia:** Model B vs A showed significant

719 improvement only for Year 3 with p-value=0.042 but not for Year 5 with p-value=0.185. Model C

720 vs A/B showed significant improvement p-values= $\sim(0.004, 0.019)$ for Year (3, 5). Model C vs D

721 showed comparable performance for Year 3 with p-value=0.156 but significant improvement for

722 Year 5 with p-value=0.025.

723 **ROW 3 NCI and incident cognitive impairment:** Model B vs A showed significant

724 improvement p-values p-values= $(0.002, 0.008)$ for Year (3, 5). Model C vs B showed

725 significant improvement p-values= $(0.001, 0.001)$ for Year (3, 5). Model C vs D showed

726 significant improvement p-values= $(0.029, 0.002)$ for Year (3, 5).

727

728

729 **Fig 2. Important predictive covariates predicting incident Alzheimer’s dementia**
730 **and incident cognitive impairment.**

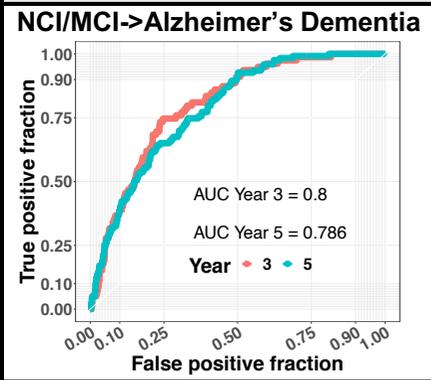
731 **ROW 1 No dementia (NCI or MCI) at baseline and incident Alzheimer’s Dementia:** Age and
732 APOE E4 allele were identified as the top positive risk factors in all three models. Hand strength
733 was the top protective factor in Model A when no cognitive covariates were considered. MMSE
734 was the top protective factor in Model B when composite cognition was not considered.
735 Composite cognition score was the top protective factor in Model C where all clinical covariates
736 were considered.

737 **ROW 2 NCI at baseline and incident Alzheimer’s Dementia:** Age was identified as the top
738 positive risk factor in all three models. Motor gait and hand strength were top protective factor in
739 Model A and Model B when composite cognition score was not considered. Instead of
740 composite cognition score, vascular disease history at baseline was the top “protective” factor in
741 Model C, which might lead to death before the development of Alzheimer’s dementia.

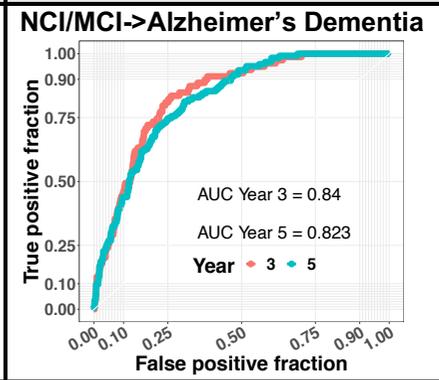
742 **ROW 3 NCI and incident cognitive impairment:** Age and education were identified as top
743 positive risk factors in all three models, where education years were positively correlated with
744 the number of follow-up years in ROS/MAP. Hand strength and social network size were top
745 protective factor in Model A when no cognitive covariates were considered. MMSE was the top
746 protective factor in Model B when composite cognition was not considered. Composite cognition
747 score was the top protective factor in Model C where all clinical covariates were considered.

748

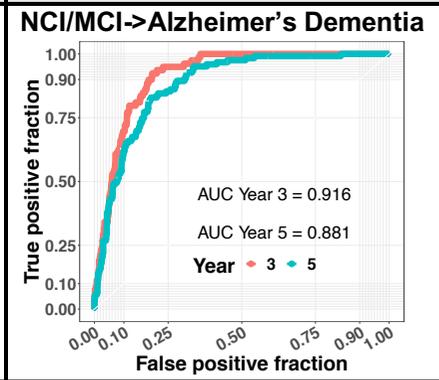
Model A. Non-Cognitive Covariates



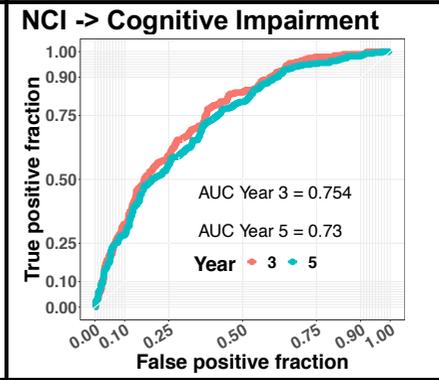
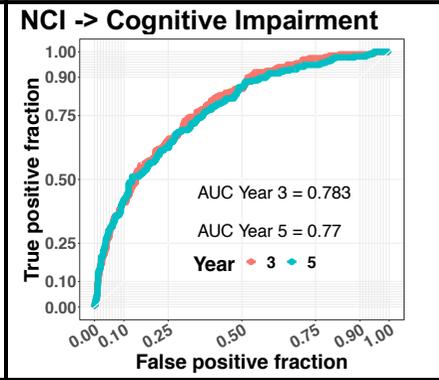
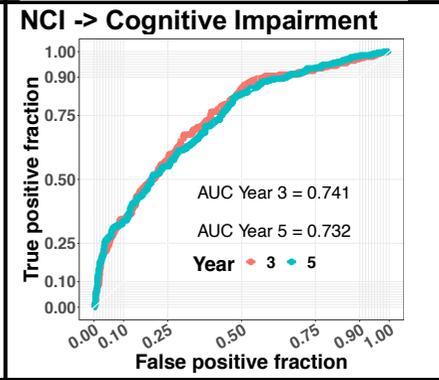
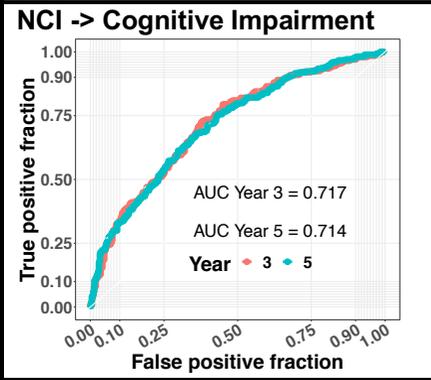
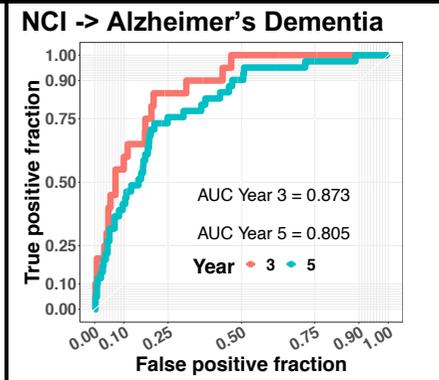
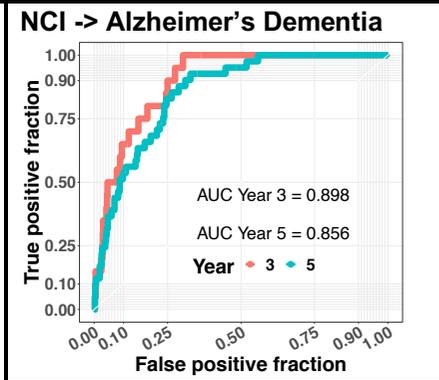
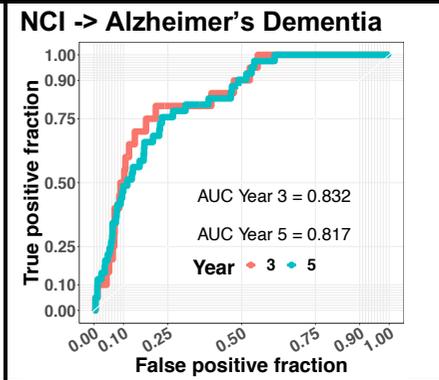
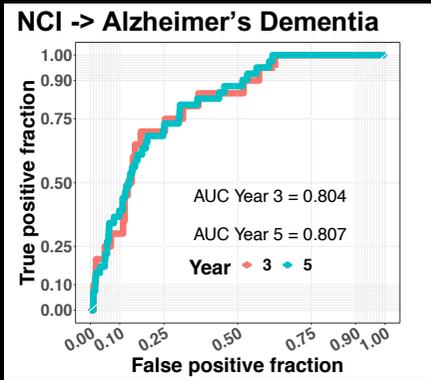
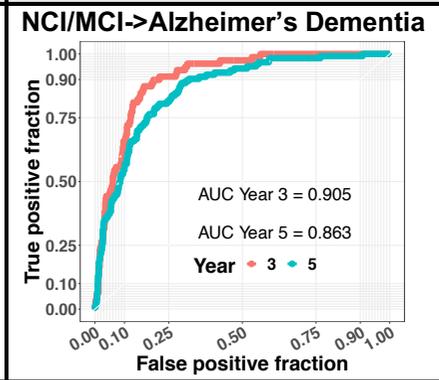
Model B. MMSE + Non-Cognitive Covariates



Model C. MMSE + Composite Cognition + Non-Cognitive Covariates

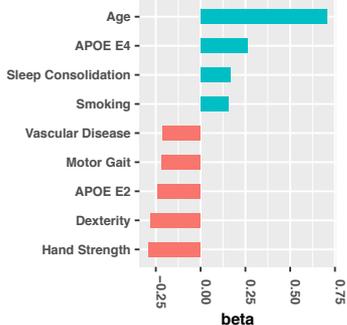


Model D. Composite Cognition Covariates



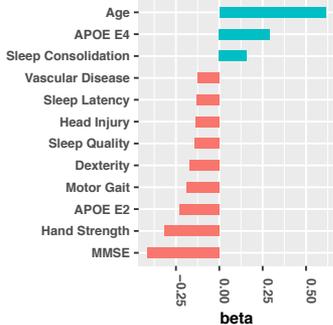
Model A. Non-cognitive Covariates

NCI/MCI->Alzheimer's Dementia



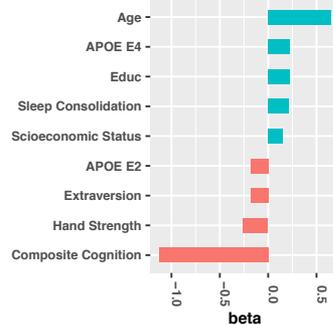
Model B. MMSE + Non-cognitive Covariates

NCI/MCI->Alzheimer's Dementia

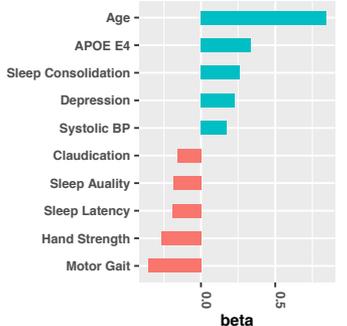


Model C. MMSE + Composite Cognition + Non-cognitive Covariates

NCI/MCI->Alzheimer's Dementia



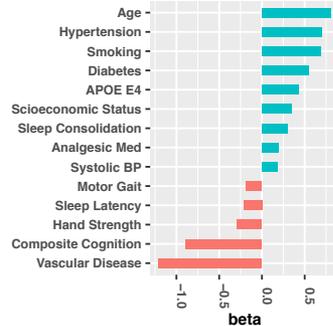
NCI -> Alzheimer's Dementia



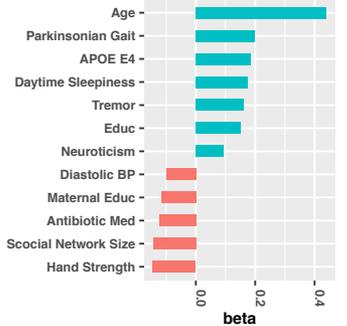
NCI -> Alzheimer's Dementia



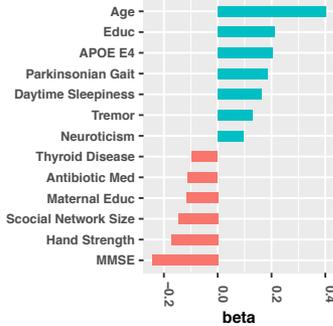
NCI -> Alzheimer's Dementia



NCI -> Cognitive Impairment



NCI -> Cognitive Impairment



NCI -> Cognitive Impairment

