ONLINE APPENDICES

Appendix A: Cognitive test descriptions

Mini-Mental State Examination (MMSE)

The MMSE is a brief mental status test measuring orientation, concentration, immediate and delayed memory, language and constructional praxis. (Folstein, et al., 1975) Scores range from 0 to 30, with higher scores indicating better cognitive performance. The MMSE has been administered in the BLSA since the mid-1980s.

Benton Visual Retention Test (BVRT)

The BVRT is a test of short-term visual memory and constructional abilities.(Benton, 1974) Administration A has been used in the BLSA since 1960, with a modified error scoring system, based on the BVRT Manual scoring, such that higher scores indicate poorer visual memory.

California Verbal Learning Test (CVLT)

The CVLT is a 16-item shopping list measuring verbal learning and memory. The variables of interest in this study were List A sum across five learning trials and long delay free recall. Scores ranged from 0 to 80 for List A sum and 0 to 16 for long delay free recall. Higher scores indicate better verbal memory. The CVLT has been administered in the BLSA since 1993, and is described in detail elsewhere. (Delis, et al., 1988)

Verbal Fluency Tests (VFT-L and VFT-C)

The verbal fluency measures of both letter (F, A, S) (Lezak, 1983,Lezak, 1995) and category (fruits, animals, vegetables) fluency (Rosen, 1980) were included. Letter fluency measures phonemic, and category fluency measures semantic fluency. Participants were required

to generate as many words as possible starting with either a specific letter or category, for 60 seconds. Higher scores indicate better verbal fluency, with the total number of words, minus intrusions and perseverations analyzed for each test. The verbal fluency tests have been administered in the BLSA since the mid 1980s.

Trails A and B:

Trail Making Tests A and B (Trails A and B) are tests of attention (Trails A) and executive functioning (Trails B), specifically cognitive control and visuo-motor scanning. (Reitan, 1992) When errors were committed the participant corrected the error by returning to his/her last correct response and continuing from there. The stop-watch recorded the time while corrections were made. Scores reflected time to completion (in seconds) separately for Trails A and B. Higher scores indicate poorer performance.

Appendix B: Linear mixed models for prediction of cognitive performance

A standard taxonomy of models (Singer and Willet, 2003) was used, starting from the unconditional means model (Model A), unconditional growth model (Model B), growth model with level-2 controlled effects of other factors namely sex, race/ethnicity, education and smoking status (Model C), growth model with level-2 controlled effects of other factors, adding a squared-age term that would allow the rate of change to vary with time (Model D). In all models, age was centered at 50 years, while education was centered at 16 years. The following equations apply to each of the models considered:

Model	Level-1 model	Level-2 model	Composite model
A	$Y_{ij} = \pi_{0i} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \zeta_{0i}$	$Y_{ij} = \gamma_{00} + (\zeta_{0i} + \varepsilon_{ij})$
B	$Y_{ij} = \pi_{0i} + \pi_{1i} Age_{50} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \gamma_{10} Age_{50} + (\zeta_{0i} + \zeta_{1i} Age_{50} + \varepsilon_{ij})$
C	$Y_{ij} = \pi_{0i} + \pi_{1i} Age_{50} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \sum_{k=1}^{7} \gamma_{0k} Z_{ik} + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \sum_{k=1}^{7} \gamma_{1k} Z_{ik} + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \sum_{k=1}^{7} \gamma_{0k} Z_{ik} + \gamma_{10} + \sum_{k=1}^{7} \gamma_{1k} Z_{ik} Age_{50} + (\zeta_{0i} + \zeta_{1i} Age_{50} + \varepsilon_{ij})$
D	$Y_{ij} = \pi_{0i} + \pi_{1i} Age_{50} + \varepsilon_{ij}$	$\pi_{0i} = \gamma_{00} + \sum_{k=1}^{7} \gamma_{0k} Z_{ik} + \zeta_{0i}$ $\pi_{1i} = \gamma_{10} + \sum_{k=1}^{8} \gamma_{1k} Z_{ik} + \zeta_{1i}$	$Y_{ij} = \gamma_{00} + \sum_{k=1}^{7} \gamma_{0k} Z_{ik} + \gamma_{10} + \sum_{k=1}^{8} \gamma_{1k} Z_{ik} Age_{50} + (\zeta_{0i} + \zeta_{1i} Age_{50} + \varepsilon_{ij})$

Notations: Y_{ij} is the response variable for each individual "i" and age at visit "j". π_{0i} is the level-1 intercept for individual i; π_{1i} is the level-1 slope for individual i; γ_{00} is the level-2 intercept of the random intercept π_{0i}; γ_{10} is the level-2 intercept of the slope π_{1i}; Z_{ik} is a vector of fixed covariates for each individual i that are used to predict level-1 intercepts and slopes; ζ_{0i} and ζ_{1i} are level-2 disturbances; ε_{ij} is the within-person level-1 disturbance.

Model D's improvement in fit compared to the simpler models was evaluated using Deviance, AIC and BIC statistics as well as pseudo-R^2. In addition, residuals were plotted against predicted values to assess their normality. It is worth noting that the models were fit using the entire BLSA cohort (n=3005) and not only those who were eligible for the main analysis to improve reliability of predicted estimates. Finally, empirical Bayes estimators of outcomes Y_{ij} were predicted from Model D at specific ages using the

following method, after estimating the random effects (ζ_{0i} for the intercept and ζ_{1i} for the slope) for each individual *i*:

2.1 $$\pi_{0i} = \gamma_{00} + \sum_{k=1}^{7} \gamma_{0k} Z_{ik} + \zeta_{0i}$$

2.2 $$\pi_{1i} = \gamma_{10} + \sum_{k=1}^{8} \gamma_{1k} Z_{ik} + \zeta_{1i}$$

2.3 $$Y_{ij} = \pi_{0i} + \pi_{1i}(Age_{50})_l$$

where $(Age_{50})_l$ is assigned individual mean age at follow-up values centered at age 50, thus positive values if Age>50 and negative values if Age<50.

Y_{ij} in this case is the cognitive score for a specific test j and individual i.