Prediction of disability-free survival in healthy older people

Johannes Tobias Neumann · Le T. P. Thao · Anne M. Murray · Emily Callander · Prudence R. Carr · Mark R. Nelson · Rory Wolfe · Robyn L. Woods · Christopher M. Reid · Raj C. Shah · Anne B. Newman · Jeff D. Williamson · Andrew M. Tonkin · John J. McNeil · on behalf of the ASPREE investigators

Abstract Prolonging survival in good health is a fundamental societal goal. However, the leading determinants of disability-free survival in healthy older people have not been well established. Data from ASPREE, a bi-national placebo-controlled trial of aspirin with 4.7 years median follow-up, was analysed. At enrolment, participants were healthy and without prior cardiovascular events, dementia or persistent physical disability. Disability-free survival outcome was defined as absence of dementia, persistent disability or death. Selection of potential predictors from amongst 25 biomedical, psychosocial and lifestyle variables including recognized geriatric risk factors, utilizing a machine-learning approach. Separate models were developed for men and women. The selected predictors were evaluated in a multivariable Cox proportional hazards model and validated internally by bootstrapping. We included 19,114 Australian and US participants aged ≥65 years (median 74 years, IQR 71.6–77.7). Common predictors of a worse prognosis in both sexes included higher age, lower Modified Mini-Mental State Examination score, lower gait speed, lower grip strength and abnormal (low or elevated) body mass index. Additional risk factors for men included current smoking, and abnormal eGFR. In women, diabetes and depression were additional predictors. The biased-corrected areas under the receiver operating characteristic
curves for the final prognostic models at 5 years were 0.72 for men and 0.75 for women. Final models showed good calibration between the observed and predicted risks. We developed a prediction model in which age, cognitive function and gait speed were the strongest predictors of disability-free survival in healthy older people.

Trial registration
Clinicaltrials.gov (NCT01038583)

Keywords Risk prediction · Disability · Survival · Elderly · Healthy · Public Health

Introduction

Currently, approximately 900 million individuals worldwide are aged 60 years or older and this number is expected to nearly double by 2050 [1]. Amongst those reaching this age in good health, future years of life will commonly be marked by functional decline involving physical disability or cognitive impairment, leading to a loss of independence and significant costs to society [2].

The ability to prolong the time spent in good health, living without dependence on others, has become a fundamental societal goal [3–5]. For this reason, disability-free survival has emerged as an important geriatric research outcome in contrast to disease-specific outcomes such as cardiovascular disease or dementia [4]. Particularly in older people, the maintenance of good health requires the avoidance of multiple interacting co-morbidities and chronic conditions, many with shared risk factors and management strategies [6–8]. Understanding the major preventable determinants, not only of individual diseases but also of ongoing good health [9], can be facilitated by development of prediction models for disability-free survival.

Previous studies have identified risk factors for a range of specific geriatric outcomes including frailty, physical disability, dementia and death. In addition to age, these have included abnormal body mass index (BMI), smoking, diabetes, abnormal blood lipids, chronic kidney disease, low level of physical activity, low gait speed, presence of depression, subclinical cardiovascular disease, specific diets and lack of social support [10–19]. These studies have typically been limited to the evaluation of pre-specified relationships between predictors and specific outcomes, rather than a composite measure of functional independence. Similarly, no previous study has employed a machine-learning approach to enable objective variable selection [20].

The ASPIrin in Reducing Events in the Elderly (ASPREE) placebo-controlled trial of low-dose aspirin was the first large-scale clinical trial that utilized disability-free survival as the prespecified primary endpoint. Measures of physical disability and cognition were assessed systematically throughout follow-up in all subjects in detail, which is rarely available in a non-trial setting. In this analysis, we used the ASPREE data to develop and validate a prediction model for disability-free survival in a population of relatively healthy individuals aged 65 or more when recruited. Machine-learning techniques allowed the relative importance of key predictors to be determined in an unbiased and objective fashion. Specifically, we sought to identify modifiable risk factors that could identify individuals at risk and prioritize preventive policies.

Methods

Study population and trial design

ASPREE was a large, randomized, double-blind, placebo-controlled trial investigating the efficacy of 100mg aspirin on extending disability-free survival in a healthy elderly population. The trial design and primary results have been published previously [21–24]. Briefly, 19,114 community-dwelling individuals without prior cardiovascular events, dementia
or physical disability were randomized to low-dose enteric-coated aspirin or placebo and followed for an average of 4.6 years. All individuals were at least 70 years old (≥ 65 years for US minorities). Exclusion criteria included a previous diagnosis of cardiovascular events (including myocardial infarction, heart failure, angina pectoris, stroke, transient ischemic attack, 50% carotid artery stenosis or previous carotid endarterectomy or stenting, coronary artery angioplasty or stenting, coronary artery bypass grafting or abdominal aortic aneurysm), atrial fibrillation, evidence of dementia or major cognitive impairment, inability to perform independently any basic Katz activity of daily living (ADL), or a serious illness with a life expectancy of less than 5 years. All participants provided written informed consent. The trial was approved by the local ethics committees and is registered on clinicaltrials.gov (NCT01038583). Requests for data access will be via the ASPREE Principal Investigators with details for applications provided through the web site, www.ASPREE.org.

Baseline collection, definition and selection of potential predictors

A total of 25 candidate predictors were selected on the basis of previous outcome studies related to dementia, disability and death [11, 17–19, 25] and their likely availability in typical clinical practice. The preselection of candidate predictors was also performed to minimize the possibility of including noise variables in the final model. The variables were collected as part of the initial standardized screening process prior to participant enrolment in ASPREE and included:

- demographics (age, living status, years of education, ethnicity/race),
- prevalent conditions or risk factors (diabetes, smoking history, alcohol consumption, family history of myocardial infarction),
- laboratory results (high-density lipoprotein cholesterol [HDL-c], non-HDL-c, estimated glomerular filtration rate [eGFR], haemoglobin),
- physical measurements (systolic and diastolic blood pressure, BMI, abdominal circumference, grip strength, gait speed),
- medication use (antihypertensive agents, lipid-lowering agents and in-trial randomization to aspirin treatment),
- measures of cognitive function (Modified Mini-Mental State Examination [3MS] score) and depression (Centre for Epidemiological Studies-Depression-10 questions [CES-D] score, as a binary variable ≥ 8).

Details related to all predictors are provided in the Supplementary Material.

Outcome ascertainment

The primary end point of disability-free survival was defined as survival free from dementia or persistent physical disability and was used as a surrogate for functional independence [21]. Physical disability and dementia were combined in the primary end point, as they represent the important reasons why individuals lose the ability to live independently. Participants without the documented outcome were censored at 5 years or at the last date they were known to be event-free, whichever came first. To assess this composite endpoint, all participants had serial in-person visits, which included standardized assessments of cognitive function with a cognitive battery, and dementia assessment if triggered, and physical disability, defined as an inability to perform, having severe difficulty in performing, or requiring assistance to complete at least one of the six basic ADL by self-report [26]. This was considered to be persistent if the same ADL loss was confirmed six months later.

Dementia was adjudicated based on the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV) criteria [27]. The ascertainment of death has been described before and was based on information collected from at least two sources including close contacts, physicians, public death notices and by linkage to the Australian National Death Index and US National Death Index [23]. Dementia and cause of death were adjudicated and confirmed by the respective endpoint committees blinded to treatment allocation.

Statistical analyses

Prognostic model development

Separate models were developed for men and women due to their known differing risk factor profiles. The statistical model employed was the Cox proportional hazards regression model. Prior to selection
of variables, we tested for non-linearity of continuous variables. Specifically, a complete case analysis was initially fitted with all candidate predictors, in which each continuous variable was modelled with a restricted cubic spline function with 2 degrees of freedom. Statistically significant non-linear relations were found for BMI, HDL-c, waist circumference and gait speed in the model for women; and non-HDL-c, eGFR, BMI and waist circumference in the model for men.

Data was missing for 11 variables with the proportion ranging from very low, <0.1% (for years of education, haemoglobin and CES-D) to as high as 8% (for family history of myocardial infarction). The total proportion of missing values was 1% but using a complete case analysis would have discarded about 15% of the sample size (see Table S1). Hence, multiple imputation was performed using chain equations, assuming data were missing at random [28]. The imputation model included all candidate predictors and other potentially relevant variables. Non-linear BMI, waist circumference, gait speed, eGFR and non-HDL-c were imputed passively [29]. The survival outcomes were included as a combination of Nelson-Aalen estimator of the cumulative hazard function and the outcome status [30]. Five data sets were imputed for men and women separately, based on the amount of missing data [31].

Variable selection was performed using the machine-learning technique “group least absolute shrinkage and selection operator” (group-lasso) in combination with bootstrapping [32]. The group-lasso model ensures that categorical variables, or linear and non-linear terms of continuous variables, are included or excluded in the model altogether. The penalty parameter, which regulates the amount of shrinkage, was selected by optimizing the 10-fold cross-validation prediction error. We chose a suboptimal penalty based on the “one-standard-error” rule to obtain a more parsimonious model [33].

The combination with bootstrap can be described as follows: for each imputed data set, we randomly drew 100 bootstrap samples with replacement, and performed group-lasso selection on each bootstrap sample. The variable inclusion frequency (VIF) over $100 \times 5 = 500$ models was calculated for each predictor. A variable with VIF $\geq 60\%$ was then included in the final model. The final model was derived by refitting the selected predictors to each of the 5 imputed datasets, and combining parameter estimates using Rubin’s rule [34]. Variable selection based on VIF has previously demonstrated good performance in the presence of imputed data [35].

Model performance

The area under the cumulative/dynamic ROC curve (AUC) at 5 years was used to assess discrimination [36]. AUC ranges from 0.5 to 1 with a higher value indicating better ability to discriminate those who developed events and those who did not. Harrell’s calibration plot at 5 years was used to assess the agreement between predicted and observed risks [37]. The apparent performances were obtained by evaluating the final model on the development samples, which were used to build the models (averaged across 5 imputed data sets).

Model validation

To quantify the degree of optimism due to overfitting in performance measures, we implemented internal validation using the enhanced bootstrap resampling procedure [37, 38]. The optimism was calculated as follows. From the original data, 100 random bootstrap samples were drawn. For each sample, we repeated the model development procedure (including imputation and lasso selection) as outlined above to obtain a bootstrap final model. We then calculated the difference between bootstrap apparent performance (averaged across bootstrap imputed data) and bootstrap test performance (averaged across original imputed data). Finally, these differences were averaged across 100 bootstrap samples to obtain the single estimate for the optimism. The procedure is illustrated in Figure S1. The estimated optimism was then subtracted from the apparent performance to obtain the bias-corrected predictive performance.

All analyses were conducted using R version 4.0.2 with the companion R packages: mice, grpreg, rms, survival, timeROC, pec.

Results

Population characteristics

A total of 19,114 community-dwelling Australian and US ASPREE participants were included in this
study. The median age of the trial population was 74 years (interquartile range [IQR] 71.6–77.7) and 10,782 (56.4%) were women. Baseline characteristics for the overall population of men and women are reported in Table 1.

During the 5-year follow-up (median [IQR] 4.6 (3.5–5.6) years), 795 (7.3%) female participants and 799 (9.6%) male participants developed dementia or persistent physical disability or died. Death accounted for 54% of the total events in men and 43% in women (Table S2).

Model development for risk of the primary composite outcome

Results of the variable inclusion frequency are shown in Table S3. The multivariable Cox regression models which included variables with VIF ≥ 60% are presented in Table 2. Specifically, increasing age, lower 3MS score, lower gait speed, lower grip strength and abnormal (low or elevated) BMI were identified as risk factors in both sexes. In men, current smoking and eGFR were additionally selected as predictors. In women, the presence of diabetes and a CES-D score ≥ 8 were also selected as predictors.

The final models (Table 2) show that a 5-year increase in age corresponded to 49% and 50% increased risk of not maintaining disability-free survival in men and women respectively, while a 5-unit increase in 3MS corresponded to about 30% decreased risk in both sexes. A decrease in gait speed from 1.0 to 0.5 m/s corresponded to 64% increased risk in men, and nearly 3-fold higher risk in women. Figure S3 illustrates the non-linear relation between gait speed and the primary outcome for women. Importantly, gait speed less than 1 m/s was associated with a substantially greater negative impact on risk of disability-free survival relative to the lesser beneficial impact of faster gait speed. The relationship between BMI and the outcome was non-linear in both models (Figure S2, S3). The risk was minimal with a BMI of around 27 kg/m², whereas both elevated and lower BMI values were associated with an unfavourable outcome.

In women, having diabetes and depression (CES-D ≥ 8) were associated with 30% and 47% increased risk of the primary composite outcome, respectively. In men, current smokers had almost 2-fold higher risk of earlier onset of death, dementia or disability compared to former/never smokers. The non-linear relation of eGFR (Figure S2) indicated that compared to a reference value of 74, a higher or lower value of eGFR was associated with a higher risk.

Model performance and calibration

The discriminative performance of each predictor and their combinations for the prediction of disability-free survival are presented in Figure 1 and Table S4. The AUCs of age alone were 0.66 and 0.65 in women and men, respectively. Adding 3MS increased the discriminative performance by about 23% in both sexes. Compared to the simplified model which consisted of age and 3MS, the final models improved the predictive performance by 24% (AUC 0.76 vs. 0.71) in women and 21% (AUC 0.73 vs. 0.69) in men (Table S4). After correcting for optimism, the final models showed good discrimination, with an AUC of 0.72 for males and 0.75 for women (Table S4).

The final models also showed good agreement between the observed and the predicted risk of the overall outcome, although the risk for men was slightly overestimated in the higher risk categories (Figure 2). The predicted probability of the endpoint at 5 years can be calculated using the formula provided in Table 3. The illustrations for two specific participants are displayed in the supplementary material (Table S5).

The risk distribution within each 5-year increment age group was further investigated in Figure S4. As expected, the figure demonstrates the strong association between advanced age and event risk in both sexes. Compared to older age groups (75–85+), younger age groups (65–74) appear to have a lower proportion of people in the lower risk categories.

Discussion

We utilized a machine-learning variable selection approach to identify the determinants of survival free of persistent physical disability and dementia in the ASPREE study population [21]. Disability-free survival is a surrogate measure of functional independence and was assessed by regular in-person surveillance of a group of 19,114 apparently healthy individuals participating in a large-scale trial of low-dose aspirin. All participants were free of prior cardiovascular events, disability and dementia at the commencement of the study.
| Characteristic                                      | All participants | Male | Female |
|----------------------------------------------------|------------------|------|--------|
| N = 19,114                                         | 74.0 (71.6, 77.7)| 73.8 (71.6, 77.3) | 74.0 (71.7, 77.9) |
| N = 8,332                                          | 8,332            |      |        |
| N = 10,782                                         |                  |      |        |
| **Age (years)**                                    |                  |      |        |
| White                                              |                  |      |        |
| Black                                               |                  |      |        |
| Hispanic                                            |                  |      |        |
| Others                                              |                  |      |        |
| **Ethnicity/race**                                  |                  |      |        |
| White                                               | 17,450 (91.3%)   |      | 9,769 (90.6%) |
| Black                                               | 901 (4.7%)       |      | 594 (5.5%)   |
| Hispanic                                            | 488 (2.6%)       |      | 286 (2.7%)   |
| Others                                              | 275 (1.4%)       |      | 133 (1.2%)   |
| **Country**                                         |                  |      |        |
| Australia                                           | 16,703 (87.4%)   |      | 9,179 (85.1%) |
| US                                                  | 2,411 (12.6%)    |      | 1,603 (14.9%) |
| **Living status**                                   |                  |      |        |
| At home alone                                       |                  |      |        |
| At home with family, friends or spouse              |                  |      |        |
| **Years of education**                              |                  |      |        |
| < 9                                                 | 3,002 (15.7%)    |      | 1,655 (15.4%) |
| 9–11                                                | 5,634 (29.5%)    |      | 3,328 (30.9%) |
| 12                                                  | 2,319 (12.1%)    |      | 1,372 (12.7%) |
| 13–15                                               | 3,255 (17.0%)    |      | 1,911 (17.7%) |
| 16                                                  | 1,766 (9.2%)     |      | 1,032 (9.6%)   |
| 17–21                                               | 3,137 (16.4%)    |      | 1,483 (13.8%) |
| **Type 2 diabetes**                                 |                  |      |        |
| Smoking history                                     |                  |      |        |
| Current                                             |                  |      |        |
| Former/never                                        |                  |      |        |
| **Alcohol consumption**                             |                  |      |        |
| Current                                             | 14,642 (76.6%)   |      | 7,710 (71.5%) |
| Former/never                                        | 4,472 (23.4%)    |      | 3,072 (28.5%) |
| Family history of myocardial infarction             |                  |      |        |
| History of cancer                                   |                  |      |        |
| Non-HDL-c (mmol/L)                                  |                  |      |        |
| HDL-c (mmol/L)                                      |                  |      |        |
| eGFR (mL/min per 1.73 m²)                           |                  |      |        |
| Haemoglobin (g/dL)                                  |                  |      |        |
| Systolic blood pressure (mmHg)                      |                  |      |        |
| Diastolic blood pressure (mmHg)                     |                  |      |        |
| Body mass index (kg/m²)                             |                  |      |        |
| Waist circumference (cm)                            |                  |      |        |
| Grip strength (kg)                                  |                  |      |        |
| Gait speed (m/s)                                    |                  |      |        |
| Lipid-lowering agents                               |                  |      |        |
| Antihypertensive agents                             |                  |      |        |
| Treatment randomization                             |                  |      |        |
| Placebo                                             | 9,589 (50.2%)    |      | 5,409 (50.2%) |
| Aspirin                                             | 9,525 (49.8%)    |      | 5,373 (49.8%) |
| 3MS                                                 | 19,114 (91.1%)   |      | 95 (92.9%)   |

Table 1 Baseline characteristics of the trial population
During a median follow-up of 4.7 years, the average age increased from 74.0 to 78.8 years and 1,594 individuals became disabled or died. By choosing a population relatively healthy at baseline and regularly monitoring each individual for the onset of dementia or physical disability, this study provided a unique opportunity to identify and prioritize the characteristics of those most likely and those least likely to continue living independently with increasing age. A particular strength of these analyses was the objective approach used for variable selection, which was based on a machine-learning technique applied to a range of clinical variables measured at study entry, together with previously recognized risk factors for physical and cognitive decline.

Out of 25 potential biomedical and social predictors tested, 7 contributed substantially to the prediction of disability-free survival in women and 7 were predictors in men. As expected, advanced age was the most important predictor, which alone resulted in an AUC of 0.65. Cognitive function was the second most important contributor followed by gait speed. Importantly, these three predictors alone resulted in the highest AUC increase, which emphasizes their importance. Additional predictors in men included grip strength, smoking, BMI and renal function. In

### Table 1 (continued)

| Characteristic | All participants | Male | Female |
|---------------|-----------------|------|--------|
|               | N = 19,114      | N = 8,332 | N = 10,782 |
| CES-D (≥8)    | 19,110          | 8,311 | 10,779 |
|               | 1,879 (9.8%)    | 631 (7.6%) | 1,248 (11.6%) |

The numbers are presented as median with interquartile range or as absolute and relative numbers. Abbreviations: N number of participants with available data, HDL-c high-density lipoprotein cholesterol, 3MS Modified Mini-Mental State, CES-D Centre for Epidemiologic Studies-Depression 10 question assessment

### Table 2 Multivariable Cox regression analyses based on the selected predictors of disability-free survival. Estimates were pooled across multiple imputed data sets using Rubin’s rule

| Characteristic | Male | | | Female | | |
|----------------|------|-----------------|-----------------|
|                | HR   | 95% CI           | p-value         | HR   | 95% CI           | p-value         |
| Age (+5 years) | 1.49 | 1.38, 1.60       | <0.001          | 1.50 | 1.40, 1.61       | <0.001          |
| 3MS (+5 units) | 0.74 | 0.69, 0.79       | <0.001          | 0.68 | 0.63, 0.72       | <0.001          |
| Gait speed (+0.2 m/s)* | 0.82 | 0.77, 0.88       | <0.001          | 0.82 | 0.77, 0.88       | <0.001          |
| 0.5 vs. 1.0 m/s | 1.64 | 1.38, 1.95       | <0.001          | 2.71 | 2.20, 3.33       | <0.001          |
| 1.5 vs. 1.0 m/s | 0.61 | 0.51, 0.73       | 0.001           | 0.80 | 0.56, 1.15       | 0.003           |
| Grip strength  | 0.92 | 0.88, 0.96       | <0.001          | 0.91 | 0.85, 0.97       | 0.003           |
| Current smoking| 1.95 | 1.51, 2.52       | <0.001          | 1.75 | 1.34, 2.27       | <0.001          |
| BMI (kg/m²)*   | 25 vs. 30 kg/m² | 1.15 | 1.05, 1.26 | <0.001 | 1.11 | 1.03, 1.19 | <0.001 |
|                | 35 vs. 30 kg/m² | 1.17 | 1.02, 1.35 | 1.29 | 1.17, 1.41 |
| eGFR (mL/min per 1.73 m²)* | 40 vs. 75 mL/min per 1.73 m² | 1.61 | 1.26, 2.07 | <0.001 | ... | ... |
|                | 90 vs. 75 mL/min per 1.73 m² | 1.24 | 1.04, 1.49 | ... | ... |
| CES-D ≥8       | ... | ...               | ...             | 1.47 | 1.23, 1.77       | <0.001          |
| Diabetes       | ... | ...               | ...             | 1.30 | 1.06, 1.60       | 0.012           |

In the model for male participants, the effect of BMI and eGFR on not maintaining disability-free survival outcome were significantly nonlinear. In the model for female participants, the effects of BMI and gait speed on the outcome were significantly nonlinear. All non-linear relations were modelled with a restricted cubic spline function with 2 degrees of freedom. To simplify interpretation of the corresponding regression coefficients, only HRs for 2 derived contrasts for each predictor are given. p values for those predictors are based on overall Wald tests of the restricted cubic spline function. Abbreviations: HR hazard ratio, CI confidence interval.
women, diabetes, and depression scores also contributed to the model. In the case of eGFR and BMI, the association was non-linear and U-shaped, similar to that noted in prior studies of mortality and other geriatric outcomes.

Previous studies have identified similar risk factor profiles as predictors of cognitive decline, frailty, functional decline and a range of other chronic diseases of ageing [10–19]. For example, a high BMI, slow gait speed and/or the presence of diabetes or hypertension has been identified as predictors of cognitive decline. Similarly slowing gait speed, weaker handgrip, increased BMI and diabetes have been identified as predictors of physical decline and frailty. The interdependence of several of these predictors creates a web of causation in which physical disability, dementia or death can be the end result of several pathways.

Figure 1 Discrimination of each selected predictor, of their combination when added sequentially in order of their AUC, and of the final models. Orange dots show the AUC at 5 years of models made with each predictor of not maintaining disability-free survival individually. Green dots show the AUC of models made by incrementally adding each predictor along the x-axis. The percentages show the added value of the current model against the previous model. Predictors are arranged by their inclusion frequency. Abbreviation: AUC, area under the curve

Figure 2 Calibration graph comparing the observed risk (based on Kaplan-Meier estimates) and the predicted risk with the final models, by tenths of the predicted probability. The red asterisk represents the bias-corrected predicted risks
The dominant impact of cognitive decline and reduced physical function suggests the possibility that interventions designed to improve cognition or physical function might alter the trajectory of decline in otherwise healthy older subjects. Recently published trials suggested that modest improvement may be possible. Amongst these, first, the LIFE study showed that the initiation of a moderate-intensity physical program in sedentary individuals aged 70–89 years modestly reduced progression to major mobility disorder, (assessed as the ability to walk 400meters). The program involving aerobic, resistance and flexibility training is provided in centre twice weekly and twice weekly at home [39]. The intervention over 2.6 years did not affect cognitive decline or reduce the incidence of cardiovascular disease [40]. Second, the SPRINT-MIND trial showed a reduction in the combined outcome of mild cognitive impairment plus probable dementia amongst older individuals on intensified blood pressure management, (assessed as the ability to walk 400metres). The program involving aerobic, resistance and flexibility training is provided in centre twice weekly and twice weekly at home [39]. The intervention over 2.6 years did not affect cognitive decline or reduce the incidence of cardiovascular disease [40]. Second, the SPRINT-MIND trial showed a reduction in the combined outcome of mild cognitive impairment plus probable dementia amongst older individuals on intensified blood pressure management [40]. The study involved 9,361 hypertensive older adults with increased cardiovascular risk whose mean systolic BP was 121.4 mmHg in the intensive-treatment group and 136.2 mmHg in the standard group. Third, the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) trial investigated the efficacy of a multidomain lifestyle intervention including dietary guidance, physical activity, cognitive training, and monitoring and management of cardiovascular risk factors. The control group received general health advice. After 2 years, the intervention group improved more in executive function, processing speed and complex memory tasks regardless of genetic risk or baseline risk factor levels.

From a clinical perspective, the identification of cognitive decline and physical limitations indicate that an individual is tracking towards an earlier loss of independence. The results of the trials described above suggest that the loss of independence in older individuals may be slowed by a multifactor intervention. However, the likelihood of these interventions producing a significant slowing of the trajectory towards loss of independence is as yet unclear.

From a population perspective, the results also emphasize the relevance of addressing the determinants of physical and cognitive decline at an earlier stage in the life course. Although diet and lipids were not included as predictive factors in the final model, their role as shared determinants in the development of cognition and physical capacity at earlier stages has been well established, and they should be included in any strategy to promote healthy aging. Furthermore, BMI and diabetes were part of the final models. A 2015 forecast of future trends in disability and life expectancy in the UK predicted an increase of 25% (over 10 years) in the number with care needs, with one quarter of the remaining life expectancy of those aged 65 years or above expected to involve a significant disability [42]. Although the increase in prevalence is largely a reflection of population ageing, it emphasizes the importance of strategies to extend disability-free survival. However, as our study results were based on an older population,
future longitudinal studies are warranted to evaluate the impact of these determinants.

Strengths of this study include the extensive data collected from a large population with relatively small amounts of missing data. The measurement of disability-free survival required data concerning the time of onset of physical disability, dementia and death, with diagnoses of dementia confirmed by a specialist adjudication panel; systematic collection of information of this type is rarely available outside the context of a clinical trial. The outcome data was accompanied by measurement of a wide range of clinically relevant and recognized geriatric potential prediction variables (biomedical and social) collected near the time of study initiation.

The limitations include that the majority of the study population were white Caucasians, most of whom lived in Australia, a high-income country with universal healthcare services. As a whole, the cohort had extensive use of preventative medications including statins (in 30%) and antihypertensive agents (in 35%), indicating a high level of preventive care. The generalizability of the study is also restricted to individuals reaching the 70-plus age group in relatively good health free of prior cardiovascular disease, dementia or significant physical limitations. It is also possible that there is residual confounding, and future research may identify new, relevant predictors. Finally, we were not able to perform an external validation, as comparable datasets are missing.

In conclusion, in a population of healthy older people, both modifiable and non-modifiable characteristics successfully identified individuals at higher risk of dying or developing dementia or physical disability within 5 years. After age, a low 3MS score was the strongest predictor followed by slow gait speed with a lesser contribution by other risk factors. Recent trials suggest that interventions aimed at slowing cognitive and physical decline are effective but additional trials will be required to demonstrate a meaningful impact in prolonging the time individuals remain functionally independent.

Acknowledgements We thank the ASPREE trial staff, participants and general practitioners.

ASPREE investigators and committees

International Steering Committee: John McNeil (Chair and Principal Investigator), Anne Murray (Co-Chair and Co-Principal Investigator), Lawrie Beilin, Andrew Chan, Jamehl Demons, Michael Ernst, Sara Espinoza, Matthew Goetz, Colin Johnston, Brenda Kirpach, Danny Liew, Karen Margolis, Frank Meyskens, Mark Nelson, Chris Reid, Raj Shah, Elsdon Storey, Andrew Tonkin, Rory Wolfe, Robyn Woods, John Zalberg

International End Point Adjudication Committees: Mark Nelson (Chair), Diane Ives (Co-Chair), Michael Berk, Wendy Bernstein, Donna Brauer, Christine Burns, Trevor Chong, Geoff Cloud, Jamehl Demons, Geoffrey Donnan, Charles Eaton, Paul Fitzgerald, Peter Gibbs, Andrew Haydon, Michael Jelinek, Finlay Macrae, Suzanne Mahady, Mobin Malik, Karen Margolis, Catriona McLean, Anne Murray, Anne Newman, Luz Rodriguez, Suzanne Satterfield (deceased), Raj Shah, Elsdon Storey, Jeanne Tie, Andrew Tonkin, Gijsberta van Londen, Stephanie Ward, Jeff Williamson, Erica Wood, John Zalberg

Data and Safety Monitoring Board: Jay Mohr (Chair), Garnet Anderson, Stuart Connolly, Larry Friedman, JoAnn Manson, Mary Sano, Sean Morrison, Erik Magnus Ohman

Australian Management Committee: John McNeil (Chair), Robyn Woods (Deputy Chair), Walter Abhayaratna, Lawrie Beilin, Geoffrey Donnan, Peter Gibbs, Colin Johnston, Danny Liew, Trevor Lockett, Mark Nelson, Chris Reid, Nigel Stocks, Elsdon Storey, Andrew Tonkin, Rory Wolfe, John Zalberg

Publications, Presentations and Ancillary Studies Committee: Anne Murray (Chair), Chris Reid (Co-Chair), Walter Abhayaratna, Michael Ernst, Colin Johnston, Beth Lewis, Danny Liew, Karen Margolis, John McNeil, Mark Nelson, Anne Newman, Thomas Obisesan, Raj Shah, Elsdon Storey, Robyn Woods

International Data Management Committee: Chris Reid (Chair), Jessica Lockery (Co-Chair), Michael Ernst, Dave Gilbertson, Brenda Kirpach, Raj Shah, Rory Wolfe, Robyn Woods

ASPREE Data Management Center (Monash University) and Biostatistics: Jessica Lockery (Data Manager), Tayla Collyer, Jason Rigby; Programmers—Kunnapoj Pruskawongsin, Nino Hay; Biostatisticians—Rory Wolfe (Senior Biostatistician), Joanne Ryan, Kim Jachno, Catherine Smith; End point Processing—A.R.M.Saifuddin Ekram (Clinical Case Reviewer), Madeleine Gardam, Henry Luong, Tim Montgomery, Megan Plate, Laura Rojas, Anna Tominaga, Katrina Wadeson

Australian Training, Recruitment, Retention and Operations Committee: Suzanne Orchard (Chair), Sharyn Fitzgerald, Sarah Hopkins, Jessica Lockery, Trisha Nichols, Ruth Trevaks, Robyn Woods

U.S. Operations/Recruitment and Retention Committee: Brenda Kirpach (Chair), Ashley Johnson, Anne Murray, Molly Prozinski, Ramona Robinson-O’Brien, Nate Tessum

Site Principal and Gp Associate Investigators:

U.S. (2,411 participants): John Aloia, Steve Anton, Jeffery Burns, Gary Burton, Jamehl Demons, Charles Eaton, Michael Ernst, Sara Espinoza, Darron Ferris, Mahalakshmi Honasoge, Daniel Hsia, Steven Katzman, Anupama Kottam, Beth Lewis, Karen Margolis, Anne Murray, Shawnah Nesbit, Anne Newman, Thomas Obisesan, Augusto Ochoa, Pricilla Pem, Kevin Peterson, James Powell, Gregg Pressman, William Robinson III, Susanne Satterfield (deceased), Raj Shah, Christine Thorburn, Elena Volpi, Jocelyn Wiggins, Jeff Williamson, Peter Wilson, Catherine Womack

Australia (GP with at least one ASPREE participant; 16,703 participants): Abdullah M. Abdul-Ridha S, Aboud E,
Funding  Open Access funding enabled and organized by CAUL and its Member Institutions This study was financially supported by the National Institute on Aging and the National Cancer Institute at the National Institutes of Health (grant numbers U01AG029824, U19AG062682); the National Health and Medical Research Council of Australia (grant numbers 334047, 1127060); Monash University; and the Victorian Cancer Agency. J.N. is recipient of a fellowship by the Deutsche Forschungsgemeinschaft (NE 2165/1-1). C.M.R. is supported through a NHMRC Principal Research Fellowship (APP 1136372). JMCN is supported through an NHMRC Leadership Fellowship (IG 1173690).

Declarations

Ethics statement  The trial was approved by the local Ethics Committees and is registered on clinicaltrials.gov (NCT01038583). All participants provided written informed consent.

Conflict of interest  The authors declare no competing interests.

Open Access  This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. World Health Organization. World report on ageing and health. 2015.

2. Disease GBD, Injury I, Prevalence C. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392(10159):1789–858.

3. Fried LP, Rowe JW. Health in aging - past, present, and future. N Engl J Med. 2020;383(14):1293–6.

4. Rowe JW, Kahn RL. Successful aging. Gerontologist. 1997;37(4):433–40.

5. Jacob ME, Yee LM, Diehr PH, Arnold AM, Thielke SM, Chaves PH, et al. Can a healthy lifestyle compress the disabled period in older adults? J Am Geriatr Soc. 2016;64(10):1952–61.

6. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet. 2012;380(9836):37–43.

7. Lafortune L, Martin S, Kelly S, Kuhn I, Remes O, Cowan A, et al. Behavioural risk factors in mid-life associated with successful ageing, disability, dementia and frailty in later life: a rapid systematic review. PLoS One. 2016;11(2):e0144405.

8. Koene RJ, Prizment AE, Blaes A, Konety SH. Shared risk factors in cardiovascular disease and cancer. Circulation. 2016;133(11):1104–14.

9. Fried LP, Guralnik JM. Disability in older adults: evidence regarding significance, etiology, and risk. J Am Geriatr Soc. 1997;45(1):92–100.

10. Elwood P, Galante J, Pickering J, Palmer S, Bayer A, Ben-Shlomo Y, et al. Healthy lifestyles reduce the incidence of chronic diseases and dementia: evidence from the Caerphilly cohort study. PLoS One. 2013;8(12):e81877.

11. Artaud F, Dugravot A, Sabia S, Singh-Manoux A, Tzourio C, Elbaz A. Unhealthy behaviours and disability in older adults: three-city Dijon cohort study. BMJ. 2013;347:f4240.

12. Newman AB, Arnold AM, Naydeck BL, Fried LP, Burke GL, Enright P, et al. “Successful aging”: effect of subclinical cardiovascular disease. Arch Intern Med. 2003;163(19):2315–22.

13. Burke GL, Arnold AM, Bild DE, Cushman M, Fried LP, Newman A, et al. Factors associated with healthy aging: the cardiovascular health study. J Am Geriatr Soc. 2001;49(3):254–62.

14. Terry DF, Pencina MJ, Vasan RS, Murabito JM, Wolf PA, Hayes MK, et al. Cardiovascular risk factors predictive for survival and morbidity-free survival in the oldest-old Framingham Heart Study participants. J Am Geriatr Soc. 2005;53(11):1944–50.

15. Willcox BJ, He Q, Chen R, Yano K, Masaki KH, Grove JS, et al. Midlife risk factors and healthy survival in men. JAMA. 2006;296(19):2343–50.

16. Yates LB, Djousse L, Kurth T, Buring JE, Gaziano JM. Exceptional longevity in men: modifiable factors associated with survival and function to age 90 years. Arch Intern Med. 2008;168(3):284–90.

17. Franzon K, Byberg L, Sjogren P, Zethelius B, Cederholm T, Kilander L. Predictors of independent aging and survival: a 16-year follow-up report in octogenarian men. J Am Geriatr Soc. 2017;65(9):1953–60.
18. Zhang S, Tomata Y, Discacciati A, Otsuka T, Sugawara Y, Tanji F, et al. Combined healthy lifestyle behaviors and disability-free survival: the Ohsaki Cohort 2006 Study. J Gen Intern Med. 2019;34(9):1724–9.
19. Bosnes I, Nordahl HM, Stordal E, Bosnes O, Myklebust TA, Almkvist O. Lifestyle predictors of successful aging: a 20-year prospective HUNT study. PLoS One. 2019;14(7):e0219200.
20. Swindell WR, Ensrud KE, Cawthon PM, Cauley JA, Cummings SR, Miller RA, et al. Indicators of “healthy aging” in older women (65–69 years of age). A data-mining approach based on prediction of long-term survival. BMC Geriatr. 2010;10:55.
21. McNeil JJ, Woods RL, Nelson MR, Reid CM, Kirpach B, Wolfe R, et al. Effect of aspirin on disability-free survival in the healthy elderly. N Engl J Med. 2018;379(16):1499–508.
22. McNeil JJ, Wolfe R, Woods RL, Tonkin AM, Donnan GA, Nelson MR, et al. Effect of aspirin on cardiovascular events and bleeding in the healthy elderly. N Engl J Med. 2018;379(16):1509–18.
23. McNeil JJ, Nelson MR, Woods RL, Lockery JE, Wolfe R, Reid CM, et al. Effect of aspirin on all-cause mortality in the healthy elderly. N Engl J Med. 2018;379(16):1519–28.
24. Group AI. Study design of ASPirin in Reducing Events in the Elderly (ASPREE): a randomized, controlled trial. Contemp Clin Trials. 2013;36(2):555–64.
25. Newman AB, Murabito JM. The epidemiology of longevity and exceptional survival. Epidemiol Rev. 2013;35:181–97.
26. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. J Am Geriatr Soc. 1983;31(12):721–7.
27. Association AP. Diagnostic and statistical manual of mental disorders. Washington, DC: DSM IV; 1994.
28. van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. J Stat Softw. 2011;45(3):1–67.
29. Stef van Buuren. Flexible imputation of missing data. Published 12 July 2018. New York. Imprint: Chapman and Hall/CRC. https://doi.org/10.1201/9780429492259.
30. White IR, Royston P. Imputing missing covariate values for the Cox model. Stat Med. 2009;28(15):1982–98.
31. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med. 2011;30(4):377–99.
32. Breheny P, Huang J. Group descent algorithms for nonconvex penalized linear and logistic regression models with grouped predictors. Stat Comput. 2015;25(2):173–87.
33. Friedman J, Hastie T, Tibshirani R. The elements of statistical learning. New York: Springer series in statistics; 2001.
34. Little RJA, Rubin DB. Statistical analysis with missing data. Hoboken, NJ: USA: John Wiley & Sons, Inc.; 2002.
35. Thao LTP, Gekoski R. A comparison of model selection methods for prediction in the presence of multiply imputed data. Biom J. 2019;61(2):343–56.
36. Blanche P, Dartigues JF, Jacqmin-Gadda H. Review and comparison of ROC curve estimators for a time-dependent outcome with marker-dependent censoring. Biom J. 2013;55(5):687–704.
37. Harrell FE. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. New York: Springer; 2015.
38. Musoro JZ, Zwinderman AH, Puhan MA, ter Riet G, Gekoski RB. Validation of prediction models based on lasso regression with multiply imputed data. BMC Med Res Methodol. 2014;14:116.
39. Pahor M, Guralnik JM, Ambrosius WT, Blair S, Bonds DE, Church TS, et al. Effect of structured physical activity on prevention of major mobility disability in older adults: the LIFE study randomized clinical trial. JAMA. 2014;311(23):2387–96.
40. LaCroix AZ, LaMonte MJ, Applegate WB. The long view of the LIFE trial and a life’s work. J Am Geriatr Soc. 2020;68(4):686–8.
41. Group SMIItSR, Williamson JD, Pajewski NM, Auchus AP, Bryan RN, Chelune G, et al. Effect of intensive vs standard blood pressure control on probable dementia: a randomized clinical trial. JAMA. 2019;321(6):553–61.
42. Guzman-Castillo M, Ahmadi-Abhari S, Bandos P, Capewell S, Steptoe A, Singh-Manoux A, et al. Forecasted trends in disability and life expectancy in England and Wales up to 2025: a modelling study. Lancet Public Health. 2017;2(7):e307–e13.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.