

Predicting incident dementia 3-8 years after brief cognitive tests in the UK Biobank prospective study of 500,000 people

Catherine M. Calvin^{1,2*}, Tim Wilkinson^{3,4}, John M Starr^{2,5†}, Cathie Sudlow^{3,4,6}, Saskia P Hagenaars^{2,7,8}, Sarah E Harris^{2,9}, Christian Schnier³, Gail Davies², Chloe Fawns-Ritchie², Catharine R Gale², John Gallacher¹, Ian J. Deary²

1 Dementias Platform UK, Department of Psychiatry, University of Oxford Warneford Hospital, Oxford OX3 7JX, UK

2 Centre for Cognitive Ageing and Cognitive Epidemiology (CCACE), Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK

3 Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, EH16 4UX, UK

4 Centre for Clinical Brain Sciences, University of Edinburgh, Chancellor's Building, 49 Little France Crescent, Edinburgh, EH16 4SB, UK

5 Alzheimer Scotland Dementia Research Centre, University of Edinburgh, 7 George Square, Edinburgh, UK. EH8 9JZ

6 UK Biobank, Cheadle, Stockport, SK3 0SA, UK

7 Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, Kings College London, Denmark Hill, London, SE5 8AF, UK

8 NIHR Biomedical Research Centre, South London and Maudsley NHS Trust, London SE5 8AF, UK

9 University of Edinburgh Centre for Genomic and Experimental Medicine and MRC Institute of Genetics and Molecular Medicine, Western General Hospital, Crewe Road, Edinburgh, UK

* Corresponding author

Phone: 01865 613103

E-mail: catherine.calvin@psych.ox.ac.uk

† Deceased 8 December 2018

Abstract

INTRODUCTION: Prospective studies reporting associations between cognitive performance and subsequent incident dementia, have been subject to attrition bias. Furthermore, the extent to which established risk factors account for such associations requires further elucidation.

METHODS: We used UK Biobank baseline cognitive data ($n \leq 488\,130$) and electronically-linked hospital inpatient and death records during three to eight-year follow-up, to estimate risk of total dementia ($n=1051$), Alzheimer's disease ($n=352$), and vascular dementia ($n=169$) according to four brief cognitive tasks, with/without adjustment for constitutional and modifiable risk factors.

RESULTS: We found associations of cognitive task performance with all-cause and cause-specific dementia ($p < 0.01$); these were not accounted for by established risk factors. Cognitive data added up to 5% to the discriminative accuracy of ROC curve models; areas under the curve ranged from 82% to 86%.

DISCUSSION: This study offers robust evidence that brief cognitive testing could be a valuable addition to dementia-prediction models.

Keywords: Alzheimer's disease; *APOE e4*; Dementia; Modifiable risk factors; Polygenic risk; cognitive performance; ROC; Vascular dementia.

1. Introduction

Cognitive performance on psychometric tests among older adults without dementia are predictive of incident all-cause dementia and/or Alzheimer's disease (AD) according to prospective cohort studies (1,2). Performance scores for global cognitive function (i.e. a composite of tests from multiple cognitive domains) discriminate between incident dementia/AD cases and non-dementia (3–6). Tests of specific cognitive domains that are most predictive include episodic memory (5,7–9), executive function (10,11), verbal fluency (7,8,11), and processing speed (11,12). Such cognitive data could therefore be usefully incorporated into dementia risk prediction models for population health monitoring, alongside more established constitutional and modifiable risk factors. A recent systematic review identified 17 population-based prospective studies reporting on the discriminative accuracy of cognitive tests for predicting incident dementia (1), all of which reported moderate values (areas under the curve (AUC) from models adjusting for age, sex and education ranged from 0.70 to 0.89), for follow-up periods spanning one to ten years. In all studies, dementia case ascertainment was by in-person clinical assessment of study participants, and therefore model estimates are subject to self-selection bias; for example, retained cohort members have on average better cognitive function than those who drop out. This issue, compounded in older-aged cohorts, can also lead to studies being severely underpowered as well as biased. Whereas case ascertainment from routinely collected electronic data is sometimes adopted in dementia epidemiological studies (13,14), it has not so far been used to address this research topic, with the exception of studies linking childhood cognitive ability to later life dementia risk (15–17). In the present study we consider the discriminative accuracy of cognitive test performance in predicting incident dementia, ascertained by linkage to electronic hospital and mortality records, in the largest study to date: the UK Biobank prospective study of 502,617 adults.

The baseline data from UK Biobank enable the study of a range of genetic and environmental exposures, to help better understand preclinical and prodromal stages of diseases (18)—particularly with its ongoing linkage to primary (i.e. general practice databases) and secondary health databases (i.e. hospital admissions and death records) (19). Investigating health trajectories in the decades leading up to dementia onset is important for understanding its complex aetiology, and validating its associated risk factors, given the early neuropathological changes that can occur decades prior to diagnosis (20). However, whereas cognitive testing with psychometric assessment promises to be a non-invasive, convenient means of indicating associated changes to cognitive function, the baseline cognitive tests administered to UK Biobank participants were particularly brief (21). We therefore assess whether performance on these tests can aid in the prediction of incident dementia during relatively short-term follow-up of this cohort. If they can, then

these tests may be sufficiently sensitive for observing cognitive change over the longer term, particularly in conjunction with an enhanced cognitive test battery at participant follow-up (<http://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=100026>).

The further intention of this paper is: (i) to assess whether any associations between baseline cognitive performance and incident dementia may be accounted for by constitutional and/or modifiable risk factors for dementia, which could inform an understanding of mechanistic pathways; and (ii) whether cognitive scores add anything to prediction models of the disease. We are aware of only one other population cohort study to date — the German Study of Aging Cognition and Dementia — that has considered preclinical cognitive performance in conjunction with both genetic and modifiable risk factors of dementia (6). The present study considers a range of constitutional and genetic factors (i.e. family history of dementia, *APOE e4* genotype, and, polygenic risk scores for AD), as well as well-validated modifiable risk factors (22), including major depressive disorder, hypertension, low physical activity, diabetes, obesity, high cholesterol, and smoking.

2. Methods

2.1 Study population

UK Biobank recruited 502 617 people aged between 40-69 years from across the UK between 2006 and 2010, to take part in a longitudinal study of health and well-being (18). During baseline assessment participants completed touchscreen cognitive tests and self-report questionnaires, and nurse interviews for physical measures and validation of self-report health-related measures. They provided blood samples for genetic analysis and were asked for consent to follow-up through linkage of their health data to their study data. Genome-wide genotyping data were analysed from blood samples of 488,363 participants. Incident health outcomes can be identified through linkage to hospital admissions and mortality datasets.

2.2 Assessment of cognitive function

Cognitive function was assessed using four computer-administered tests (see Table 1 for details). The bespoke cognitive test battery was intentionally brief and designed to tap into cognitive domains that are sensitive to ageing and/or pathological processes (i.e. fluid cognitive abilities), including processing speed (reaction time task), visual episodic memory (visual pairs memory task), and prospective memory (prospective memory task). In addition, a verbal-

numerical reasoning task that assesses both fluid and crystallised abilities, showed the highest loading on a general cognitive ability factor according to a psychometric analysis of this battery (21). The prospective memory and verbal-numerical reasoning tasks were introduced during the last two years of UK Biobank recruitment, and therefore data on these are available for approximately one third of the cohort.

[Table 1 about here]

2.3 Ascertainment of incident dementia

Dementia syndromes, including all-cause, AD and vascular dementia (VaD), were identified from ICD9 and ICD10 (international classification of diseases, 9th and 10th revisions) codes of hospital inpatient admissions data from 1996 to February 2016 (provided by Hospital Episode Statistics in England; for corresponding start dates and annual censoring dates from Scottish Morbidity Records and Patient Episode Database for Wales see: https://biobank.ctsuo.ox.ac.uk/crystal/exinfo.cgi?src=Data_providers_and_dates), and from ICD10 codes on death certificates issued between April 2006 and February 2016 (see Table S1 for ICD codes), downloaded from UK Biobank's data showcase on 29th November 2017 (<http://biobank.ndph.ox.ac.uk/showcase/>). Data linkage processes including permissions were managed by UK Biobank, and participant identifiers (NHS number, date of birth, sex, and postcode) were submitted to external organisations for secure matching to their respective secondary healthcare records (for more details see: <https://biobank.ctsuo.ox.ac.uk/crystal/docs/DataLinkageProcess.pdf>), which included primary and secondary diagnoses within hospital admissions, and underlying and secondary causes of death from morbidity records. We selected incident dementia cases if the date of first recorded diagnosis for an individual occurred at least three years after the date of their baseline assessment, excluding participants with diagnoses earlier than this to reduce the likelihood of including prevalent cases in our analyses. We also excluded those who self-reported having 'dementia, Alzheimer's disease, or cognitive impairment' at the baseline nurse interview (n=56 not detected by hospital admissions records).

2.4 Constitutional and modifiable risk factors

Table 2 includes details of the dementia risk factors included as covariates in the present study. Constitutional risk factors for dementia include: age, sex, education, family history of dementia, *APOE e4* carrier status, AD polygenic score, and, neuroticism. We selected seven modifiable risk factors for dementia validated in a recent systemic review

and expert panel study (23), including depression (current risk and past history), hypertension, hyperlipidaemia, diabetes, obesity (and central obesity), smoking, and, low physical activity.

[Table 2 about here]

2.5 Statistical analysis

Age and sex-adjusted logistic regression models were used to represent the fundamental associations between each independent variable (cognitive, constitutional, and modifiable risk factors) and dementia risk (total, AD and VaD). The independent variables of significance in univariate analysis were then included in multiple logistic regression models (complete cases only): Model 1 included age, sex, education, and one of the four cognitive test variables; Model 2 added constitutional risk factors (i.e. *APOE e4*, family history of dementia and neuroticism) to Model 1; Model 3 added modifiable risk factors to those variables entered into Model 2, including current and past depression, hypertension, hypercholesterolemia, diabetes, and smoking. To address bias due to data not missing completely at random, these multiple logistic regression models were repeated to include participants with missing data on covariates with >5% missingness in the total sample, and results were compared with those from models including complete cases only (see Supplementary Methods B for details of missingness). Thirdly, we considered the discriminative accuracy for predicting incident dementia risk according to those variables that were shown to independently predict the outcomes, by plotting receiver operating characteristic (ROC) curves and deriving AUC.

3. Results

3.1 Descriptive characteristics of the sample

We identified 1051 cases of incident dementia (including 352 AD; 169 VaD) via electronic linkage to hospital admission and death certificate data during the period three to eight years after cognitive testing (5.2 median years of follow-up; 2.1 cases per 1000 population – slightly lower than rates reported in similarly aged UK cohorts (23,24)). Of the original UK Biobank cohort, 385 336 participants including 591 dementia cases had complete data on reaction time, visual memory, and covariates, and were therefore included in the first set of univariate analyses (‘main analytic sample’). Models including genetic data were based upon a subset of this sample: 65 824 participants and 112 dementia cases (‘genetic sample’). A second subset were used in analyses of the remaining two cognitive tests (‘full cognitive

battery sample'), involving 130 212 participants and 115 dementia cases who had complete data on verbal-numerical reasoning and prospective memory. The mean lag times between cognitive testing and dementia diagnosis were similar for the total sample ($n = 1051$; $M = 5.23$ years) and the main analytic sample ($n = 591$; $M = 5.24$ years) (see Figure S1 for distribution of lag times). Those who were excluded from analyses due to missing data were on average older at dementia diagnosis (69.4 years vs. 66.2 years), and, showed a higher dementia incidence (0.39% vs. 0.15%); they also showed slightly greater modifiable risk factors for dementia than the analytic sample, they were less likely to have graduated from college, and they were poorer scorers on average on the cognitive tests (see S5 Table). There were no differences between the analytic and excluded samples on genetic status, including *APOE e4*.

3.2 Association of individual risk factors for dementia

In age- and sex-adjusted models each constitutional risk factor—with the exception of AD polygenic score—was significantly associated with total incident dementia (see Table 3 for full results): positive *APOE e4* carrier status ($OR = 2.99$; 95%CI: 2.43 to 3.68), having a family history of dementia (1.65 [1.36 to 2.00]), no college degree (1.48 [1.22 to 1.81]), high neuroticism (1.45 [1.15 to 1.84]), male sex (1.39 [1.18 to 1.63]), , and older age (1.21 [1.19 to 1.23] per year).

Similarly, the majority of modifiable risk factors were significantly associated with dementia risk: baseline depression rating ($OR = 2.25$; 95%CI: 1.67 to 3.03), diabetes (1.75 [1.36 to 2.26]), current smoking (1.53 [1.17 to 2.01]), history of depression (1.47 [1.24 to 1.74]), hypertension (1.39 [1.17 to 1.64]), and high cholesterol (1.36 [1.14 to 1.61]). Total obesity risk (i.e. $BMI \geq 30$) and central obesity were not significantly associated with all-cause dementia risk in this population (1.12 [0.97 to 1.28]; 1.16 [0.97, 1.40]); neither was low physical activity (1.23 [0.98 to 1.54]). In sensitivity analysis the odds ratios for total dementia risk according to these modifiable risk factors were somewhat weaker compared to analyses that included incomplete cases (data available in S6 Table). In further sensitivity analyses univariate models were repeated with death included as an outcome, and survival bias was viewed as unlikely for most factors (see Table S7).

The above risk factors were differentially related to risk of AD and VaD. For example, *APOE e4* genotype was more strongly predictive of incident AD than incident VaD (4.28 [2.97 to 6.18] vs. 2.22 [1.33 to 3.70]), whereas being male was more predictive of VaD versus AD (1.81 [1.19 to 2.75] vs. 1.04 [0.78 to 1.38]). In addition, of the modifiable risk factors we investigated current depression, hypertension, high cholesterol, and diabetes, were each significantly

predictive of VaD, whereas only current depression was predictive of incident AD (see Table 3 for individual effect sizes).

[Table 3 about here]

3.3 Prediction of incident dementia by baseline cognitive performance

3.3.1 Univariate analyses

Baseline performances on all four cognitive tests were significantly predictive of incident dementia during three to eight years of follow-up, after adjustment for age, sex and education (see Table 4). A one standard deviation (SD) higher verbal-numerical reasoning error score was associated with 68% higher odds of dementia diagnosis (OR = 1.68; 95%CI: 1.36 to 2.07); a one SD higher mean reaction time (i.e., a slower reaction time, indicating poorer performance) was associated with 31% higher odds (1.31 [1.22 to 1.41]); and a one SD higher number of errors on the visual memory task was associated with 27% higher odds of dementia (1.27 [1.18 to 1.36]). Whereas the magnitude of these effect sizes are relatively low (if we compare them to those of established risk factors, i.e. *APOE e4*, current depression), an incorrect first response on the prospective memory task was associated with over three times the odds of dementia (3.28 [2.26 to 4.75]). The effect sizes for total dementia risk according to reaction time and visual memory respectively were similar for the main analytic sample versus the sample with incomplete cases (see Table S8 for sensitivity analyses). Furthermore, there were no obvious differences between the odds ratios for dementia diagnoses made within three to five years of cognitive testing, and those for dementia diagnoses made six to eight years later (see Table S8).

For the dementia subtype prediction models, reaction time was more highly predictive of VaD (1.56 [1.32 to 1.84]) than AD (1.24 [1.09 to 1.41]) (comparison of natural logs of OR: $Z=2.15$; $p=0.031$), whereas visual memory performance was associated to a similar extent with AD and VaD respectively (1.24 [1.10 to 1.41]) vs 1.20 [1.00 to 1.44]) (see Table 4). 2.86 [1.20 to 6.80]) Although models for the other two cognitive tasks showed higher point estimates for AD than VaD, confidence intervals were wide and overlapping, and case numbers low.

[Table 4 about here]

3.3.2 Adjustment for constitutional and modifiable risk factors

Despite the differential association of risk factors with incident AD and VaD respectively, we first conducted an adjusted analyses by these risk factors on the association between cognitive performance and all-cause dementia. This was warranted given the higher proportion of ‘unspecified dementia’ diagnoses in this sample compared to either subtype (Table S1), the likelihood of additional AD and VaD cases among this unspecified group, and, the subsequent loss of statistical power if we were to restrict analyses to AD and VaD endpoints only. In multiple logistic regression models sample sizes increased after excluding covariates that were non-significant in univariate analyses. Table 5 reports the results of these models where the associations between cognitive domain-related tests and all-cause dementia risk, were adjusted for constitutional risk factors (with and without *APOE e4*), then additionally for modifiable risk factors. These did little to account for any of the associations between cognitive task performance and dementia risk in the present study, which remained statistically significant in fully-adjusted models. For example, the ORs for dementia in basic and fully-adjusted models were: 1.31 (95%CI: 1.20, 1.43) and 1.31 (1.20, 1.43) according to a one SD slower reaction time; 1.21 (1.11, 1.32) and 1.19 (1.09, 1.30) for one SDs higher in visual memory errors; 1.87 (1.42, 2.48) and 1.72 (1.06, 2.81) for one SDs higher in verbal-numerical reasoning errors; 3.50 (2.15, 5.67) and 3.00 (1.30, 6.96) for an incorrect first response in prospective memory. Other independent variables that retained significance in multiple regression included age, sex, *APOE e4*, family history of AD, and self-rated depression (see Tables S9-S12 for effect estimates). Sensitivity analyses found no evidence that excluding cases due to substantive missing data on exposures would have biased these results (see Table S13).

Multiple logistic regression models were run to predict dementia subtypes in association with reaction time and visual memory respectively. In a basic model the OR for incident AD was 1.29 (95%CI: 1.12, 1.49) according to a one SD slower reaction time, compared to 1.30 (1.13, 1.50) in a fully-adjusted model (see Table S9). The ORs for VaD according to a one SD slower reaction time in the basic and fully-adjusted models were 1.58 (1.36, 1.84) and 1.53 (1.32, 1.78) respectively. Similarly, the inclusion of constitutional and modifiable risk factors did very little to account for the associations between visual memory, and AD and VaD respectively (Table S10 reports these results).

[Table 5 about here]

3.4 Discriminative accuracy of models, with and without cognitive variables

Figure 1 displays ROC curves for all-cause dementia prediction models, with age, sex and education entered in the first model, followed by the addition of family history with or without *APOE e4* (model 2), + self-rated depression

(model 3), and + cognitive performance on all four tests (model 4). Whereas genetic testing is becoming more cost-effective, the decision to determine the predictive accuracy of a model without *APOE e4* was because participants of research studies don't always consent to use of their genetic data. Therefore, the analyses included: 81 823 (78 dementia cases) with no missing data when *APOE e4* status was included (Figure 1, left graph), and, 145 068 participants (141 dementia cases) when *APOE e4* was excluded (Figure 1, right graph). The AUC values for model 4 were moderately good (83% to 86% respectively), and each of the cognitive tests contributed to the models (see Table S14). The addition of reaction time, visual memory, verbal-numerical reasoning, and prospective memory to ROC curves that included age, sex, education, family history of AD, and depression, with or without *APOE e4* status, significantly improved the models' predictive power according to chi-square tests, contributing an additional 5% to the discriminative accuracy of the ROC curves.

For comparison with previous studies we reproduced ROC curves adding the four cognitive test variables to a basic model that included age, sex and education only (n=160 903; cases=177). The AUC value went from 0.763 (95%CI: 0.729, 0.796) to 0.821 (0.790, 0.853), $\chi^2 = 31.9$, $p < .001$.

[Figure 1 about here]

Discussion

Our study has demonstrated an association between preclinical cognitive capability in adulthood and risk of incident dementia ascertained through data linkage to routine electronic health records, thereby avoiding the issue of attrition common to previous studies. In a dementia-free sample at baseline we observed associations between performance on very brief tests of reaction time, visual memory, prospective memory, and verbal-numerical reasoning and incident all-cause dementia during three to eight-year prospective follow-up. The cognitive scores were also each related to specific dementia subtypes AD and VaD. Despite including established risk factors for dementia in the models, the findings regarding preclinical cognitive performance were independent of genetic and other constitutional factors, as well as seven modifiable risk factors. The cognitive tests together added up to 5% to the discriminative accuracy of ROC curves that included age, sex, family history of dementia, *APOE e4*, and, depression (AUC was 0.86 in the fully-adjusted model).

A limitation of the present study is the potential under-detection of incident dementia cases due to referral bias. This may have been a factor in the hospital admissions data due to individual differences in health care-seeking behaviour. However, such bias may be less than expected among UK Biobank participants who are more likely to seek health care given their willingness to voluntarily participate in health research. In 2012, hospitals in England saw the introduction of routine dementia assessment for old-aged patients, which may further reduce this bias. Since then improvements in sensitivity and specificity of dementia diagnoses in hospitals in England have been reported (25), although the validity of recording VaD in hospital inpatient records, in particular, could still be improved (26,27). Identification of dementia cases through electronic health linkage without use of primary care data (unavailable to us), may also have led to under-detection of dementia, which has been observed in a Scottish subsample of UK Biobank (27). Evidence from the Million Women Study has demonstrated the reliability of using secondary health care records for dementia ascertainment, and their agreement with primary care diagnoses (28), yet, dementia codes appeared in primary care data on average 1.6 years before hospital admission data, for people who receive both. Referral bias in the present analysis is therefore likely to be inversely related to duration of follow-up. We optimised use of the secondary health care data available to us by selecting ICD codes of specific dementia subtypes as well as general dementia, and, by combining data from hospital and mortality records, which have been shown to improve positive predictive validity and sensitivity respectively of detecting dementia (14). It is uncertain how a degree of under-detection could have influenced our results, although we speculate that this may have produced more conservative odds ratios for dementia risk, at least according to the modifiable and cognitive exposures. A further limitation of the study was the low to moderate test-retest reliability of the cognitive tests (21). The UK Biobank baseline cognitive assessment was designed to be very brief due to the wide array of data types that were collected on participants, and therefore the tests were necessarily bespoke. If time had allowed for a more extensive test battery, we may have observed higher discriminative accuracy from cognitive assessment. Thirdly, missing data from participants with lower cognitive ability and/or poorer health, were likely to have reduced the sensitivity of the analysis due to selective attrition, albeit in repeat analyses the inclusion of participants with missing data saw little change to the effect sizes. Finally, the prediction modelling of the present study does not allow us to comment on causal or explanatory mechanisms for the observed associations between risk factors and incident dementia. By including education in the models, we will have controlled, to some extent, for pre-morbid cognitive ability. Yet the lower cognitive performance — as well as higher rates of depression (23) — of participants with dementia are still likely to be among the symptoms of a disease that has a long prodromal period. Still

we did not find evidence that cognitive performance was more predictive of dementia for shorter versus longer latency periods, in contrast to previous findings (3,29,30).

The strengths of the present study include the large scale of the UK Biobank cohort combined with the rich dataset that allowed us to consider multiple validated risk factors for dementia alongside the cognitive data, for prediction of all-cause and cause-specific dementia. The age range of the cohort covering mid to late adulthood also enabled us to consider cognitive function and dementia risk factors in both earlier and later onset dementias, when previous studies have recruited adults at a minimum age of 65 years (3–9,11,12,29–34). A major strength was the use of routinely-collected health records to identify dementia cases. Despite the potential issue of under-detecting cases of dementia outlined above, this method allowed us to address to a large extent the issue of attrition in previous studies that have relied upon active participation at follow-up for ascertainment of dementia, and likely lose representation from poor cognitive performers (35).

Previous longitudinal cohort studies of older adults investigating incident dementia risk according to cognitive test performance, report AUC values in the range 0.49 to 0.92 (1). Our own results from UK Biobank therefore fall into the high end of this range. For greater comparability however, if we consider our results from ROC curves with age, sex, education and cognitive test score entered (AUC: 0.82) with those studies of a similar follow-up period (three to eight years) and range of variables, we start to reach similar results: 0.73 to 0.89 (6,7,29,31,32), despite study design and selection bias differences. Perhaps what the present study lost in sensitivity from a brief cognitive assessment, it gained in sensitivity through completeness of follow-up by means of electronic linkage. Nevertheless, what the data from the present study show is that even a brief cognitive test battery has predictive power at the population level. Whereas it was not the aim of the present study to find evidence for clinical utility of such a tool, such tests do have the advantage of providing greater variation in cognitive ability scores within non-clinical population samples, than say memory clinic assessments designed to detect impairment within individuals (i.e. Mini-Mental State Examination). A brief cognitive testing tool, developed to encompass the array of cognitive domains typically affected in preclinical stages of dementia, and to be age-specific (8), could, for example, be a relatively simple, cost-effective way of recruiting an enriched cohort into dementia intervention studies (i.e. clinical trials). Whereas the UK Biobank battery included some such domains (e.g. processing speed/reaction time, visual episodic memory), it did not include others (i.e. executive function and verbal fluency) (7,8,11,12), unlike the enhanced battery that is being used for follow-up of UK Biobank participants returning for brain imaging. In future studies these enhanced cognitive data promise to be valuably incorporated into prediction modelling of later incident dementia in this cohort.

One previous study controlled for *APOE e4* status, family history of dementia, and modifiable risk factors for dementia (including depression, smoking) in its investigation of preclinical cognitive performance and dementia prediction (6) – specifically AD. Although the study did not consider whether these established risk factors accounted for any of the associations between cognitive performance and dementia risk, the authors reported on models in which scores on verbal fluency and verbal memory retained statistical significance in the presence of genetic and modifiable risk factor effects. In their most predictive of models the AUC was 0.85 (0.80 to 0.89), which compares to our own of 0.86 (0.81 to 0.90). We might have expected the effect of cognitive performance on dementia risk to reduce in multiple regression models that included genetic and modifiable (vascular) risk factors. Whereas we observed that AD risk was defined by a relatively stronger association with *APOE e4*, and VaD risk was defined more by slower processing speed and greater vascular risk factors, it was not evident that these genetic and modifiable influences accounted for the observed cognitive deficits. Hence, our data confirm that cognitive profiling adds power to prediction models of dementia, though the incremental effects are relatively limited compared with other risk factors such as age, at least within the age range of the UK Biobank sample.

A final point on the role of modifiable risk factors, with respect to the combined and subtype dementia analyses presented in Table 3, is that we acknowledge some implications of the differences in apparent aetiologic pathways between AD and VaD. None of the modifiable risk factors (with the exception of depression, which one might argue is not modifiable) increased the risk of AD, whereas several increased the risk of VaD. Genetic factors, along with increasing age, were the only other significant predictors of incident AD. Because of recent failures of anti-amyloid therapies, attention is turning to prevention of cognitive impairment via modifiable risk factor interventions. We note, in the present analysis, the fact that modifiable risk factors had almost no impact on the performance of the prediction models that included cognitive and genetic factors. When AD and VaD were combined into a single diagnosis their very different risk factor profiles tended to become obscured, and the risk associated with vascular risk factors and education seemed to acquire more salience. This was evident despite the fact that AD risk factors might have dominated the effect estimates due to the relatively higher population prevalence of AD (i.e. two-thirds). On the basis of these findings one should not lose sight of the fact that interventions to reduce dementia risk in the population may need to be very different for the dementia subtypes. Therefore future studies of this kind should attempt to more accurately characterise the subtypes, while not forgetting their greater concurrence in increasing older age.

The present study validates the use of the UK Biobank baseline cognitive data for use in dementia-related research of this cohort, and highlights that brief, electronically-delivered cognitive tests can add unique value to dementia prediction models, with potential benefits to clinical trial selection, and, to population health monitoring.

Data access and ethics

The third-party data that support the findings of this study are available from UK Biobank (Research Ethics Committee approval number 11/NW/0382), which is an open access resource for international bona fide researchers (<http://www.ukbiobank.ac.uk/register-apply/>). The authors do not have permission to distribute the data themselves.

Acknowledgements

This research has been conducted using the UK Biobank Resource under Application 10279 (<http://www.ukbiobank.ac.uk/2015/06/dr-catherine-gale-university-of-edinburgh/>).

Conflicts

J.G was the recipient of an MRC grant for staff, which included a salary for C.M.C and C.F-R; with this exception the authors declare no other financial relationships with any organisations that might have an interest in the submitted work. I.J.D is a participant in UK Biobank. C.L.M.S. is Chief Scientist for UK Biobank.

Funding Sources

This work was supported by the Medical Research Council (MRC)-funded Dementias Platform UK (MR/L023784/2), which included a salary for C.M.C, C.F-R and (for one year) T.W. T.W. also received salary funding during this study from an MRC Clinical Research Training Fellowship (MR/P001823/1). The salaries of C.F-R, S.H and G.D, and part of I.J.D's salary were funded by the Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross

council Lifelong Health and Wellbeing Initiative (MR/K026992/1); funding from the Biotechnology and Biological Sciences Research Council and MRC is gratefully acknowledged. The salary for C.L.M.S. was from UK Biobank and the Scottish Funding Council. This study represents independent research part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

1. Hou X-H, Feng L, Zhang C, Cao X-P, Tan L, Yu J-T. Models for predicting risk of dementia: a systematic review. *J Neurol Neurosurg Psychiatry* [Internet]. 2018;jnnp-2018-318212. Available from: <http://jnnp.bmj.com/lookup/doi/10.1136/jnnp-2018-318212>
2. Bäckmann L, Jones S, Berger A-K, Laukka EJ, Small BJ. Multiple cognitive deficits during the transition to Alzheimer ' s disease. *J Intern Med*. 2004;195–204.
3. Chary E, Amieva H, Pérès K, Orgogozo JM, Dartigues JF, Jacqmin-Gadda H. Short-versus long-term prediction of dementia among subjects with low and high educational levels. *Alzheimer's Dement*. 2013;9(5):562–71.
4. Wolfgruber S, Jessen F, Wiese B, Stein J, Bickel H, Mösch E, et al. The CERAD neuropsychological assessment battery total score detects and predicts alzheimer disease dementia with high diagnostic accuracy. *Am J Geriatr Psychiatry*. 2014;22(10):1017–28.
5. Restaino M, Matthews FE, Minett T, Albanese E, Brayne C, Stephan BC. Predicting risk of 2-year incident dementia using the CAMCOG total and subscale scores. *Age Ageing*. 2013;42(5):649–53.
6. Jessen F, Wiese B, Bickel H, Eiffländer-Gorfer S, Fuchs A, Kaduszkiewicz H, et al. Prediction of dementia in primary care patients. *PLoS One*. 2011;6(2).
7. Jorm AF, Masaki KH, Petrovitch H, Ross GW, White LR. Cognitive deficits 3 to 6 years before dementia onset in a population sample: The Honolulu-Asia Aging Study. *J Am Geriatr Soc*. 2005;53(3):452–5.
8. Jungwirth S, Zehetmayer S, Bauer P, Weissgram S, Tragl KH, Fischer P. Prediction of Alzheimer dementia with short neuropsychological instruments. *J Neural Transm*. 2009;116(11):1513–21.
9. Derby CA, Burns LC, Wang C, Katz MJ, Zimmerman ME, L'italien G, et al. Screening for predementia AD: Time-dependent operating characteristics of episodic memory tests. *Neurology*. 2013;80(14):1307–14.

10. Rapp MA, Reischies FM. Attention and executive control predict Alzheimer disease in late life: Results from the Berlin Aging Study (BASE). *Am J Geriatr Psychiatry* [Internet]. 2005;13(2):134–41. Available from: <http://dx.doi.org/10.1097/00019442-200502000-00007>
11. Madureira S, Verdelho A, Moleiro C, Ferro JM, Erkinjuntti T, Jokinen H, et al. Neuropsychological predictors of dementia in a three-year follow-up period: Data from the ladis study on behalf of the ladis study. *Dement Geriatr Cogn Disord*. 2010;29(4):325–34.
12. Rapp MA, Reischies FM. Attention and executive control predict Alzheimer disease in late life: Results from the Berlin Aging Study (BASE). *Am J Geriatr Psychiatry*. 2005;13(2):134–41.
13. Sibbett RA, Russ TC, Deary IJ, Starr JM. Dementia ascertainment using existing data in UK longitudinal and cohort studies: A systematic review of methodology. *BMC Psychiatry*. 2017;17(1):1–16.
14. Wilkinson T, Ly A, Schnier C, Rannikmäe K, Bush K, Brayne C, et al. Identifying dementia cases with routinely collected health data: A systematic review. *Alzheimer's Dement*. 2018;(April):1–14.
15. Russ TC, Hannah J, Batty GD, Booth CC, Deary IJ, Starr JM. Childhood Cognitive Ability and Incident Dementia. *Epidemiology* [Internet]. 2017;28(3):361–4. Available from: <http://insights.ovid.com/crossref?an=00001648-201705000-00009>
16. Calvin CM, Batty GD, Der G, Brett CE, Taylor A, Pattie A, et al. Childhood intelligence in relation to major causes of death in 68 year follow-up: Prospective population study. *BMJ*. 2017;357.
17. Whalley LJ, Starr JM, Athawes R, Hunter D, Pattie a, Deary IJ. Childhood mental ability and dementia. *Neurology*. 2000;55:1455–9.
18. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med*. 2015;12(3):1–10.
19. Allen N, Sudlow C, Downey P, Peakman T, Danesh J, Elliott P, et al. UK Biobank: Current status and what it means for epidemiology. *Heal Policy Technol* [Internet]. 2012;1(3):123–6. Available from: <http://dx.doi.org/10.1016/j.hlpt.2012.07.003>
20. Clifford RJJ, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, et al. Update on hypothetical model of Alzheimer's disease biomarkers. *Lancet Neurol*. 2013;12(2):207–16.
21. Lyall DM, Cullen B, Allerhand M, Smith DJ, Mackay D, Evans J, et al. Cognitive test scores in UK Biobank: Data reduction in 480,416 participants and longitudinal stability in 20,346 participants. *PLoS One*.

2016;11(4):1–10.

22. Deckers K, Van Boxtel MPJ, Schiepers OJG, De Vugt M, Sánchez JLM, Anstey KJ, et al. Target risk factors for dementia prevention: A systematic review and Delphi consensus study on the evidence from observational studies. *Int J Geriatr Psychiatry*. 2015;30(3):234–46.
23. Singh-Manoux A, Dugravot A, Fournier A, Abell J, Ebmeier K, Kivimäki M, et al. Trajectories of Depressive Symptoms Before Diagnosis of Dementia. *JAMA Psychiatry* [Internet]. 2017;1–8. Available from: <http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/jamapsychiatry.2017.0660>
24. Ahmadi-Abhari S, Guzman-Castillo M, Bandosz P, Shipley MJ, Muniz-Terrera G, Singh-Manoux A, et al. Temporal trend in dementia incidence since 2002 and projections for prevalence in England and Wales to 2040 : modelling study. 2017;2856(July):1–15. Available from: <http://dx.doi.org/doi:10.1136/bmj.j2856>
25. Sommerlad A, Perera G, Singh-Manoux A, Lewis G, Stewart R, Livingston G. Accuracy of general hospital dementia diagnoses in England: Sensitivity, specificity, and predictors of diagnostic accuracy 2008–2016. *Alzheimer's Dement*. 2018;14(7):933–43.
26. Thomas BM, Starr JM, Whalley LJ. Death certification in treated cases of presenile Alzheimer's disease and vascular dementia in Scotland. *Age Ageing*. 1997;26(5):401–6.
27. Wilkinson T, Schnier C, Bush K, Rannikmäe K, Henshall DE, Lerpiniere C, et al. Identifying dementia outcomes in UK Biobank : a validation study of primary care , hospital admissions and mortality data. *Eur J Epidemiol* [Internet]. 2019;(0123456789). Available from: <https://doi.org/10.1007/s10654-019-00499-1>
28. Brown A, Kirichek O, Balkwill A, Reeves G, Beral V, Sudlow C, et al. Comparison of dementia recorded in routinely collected hospital admission data in England with dementia recorded in primary care. *Emerg Themes Epidemiol* [Internet]. 2016;13(1):11. Available from: <http://ete-online.biomedcentral.com/articles/10.1186/s12982-016-0053-z>
29. MC T, Yao C, Kiss A, McDowell I. Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. *Neurology*. 2005;64(11):1853–9.
30. Tierney MC, Moineddin R, McDowell I. Prediction of all-cause dementia using neuropsychological tests within 10 and 5 years of diagnosis in a community-based sample. *J Alzheimer's Dis*. 2010;22(4):1231–40.
31. Nakata E, Kasai M, Kasuya M, Akanuma K, Meguro M, Ishii H, et al. Combined memory and executive function tests can screen mild cognitive impairment and converters to dementia in a community: The Osaki-Tajiri Project. *Neuroepidemiology*. 2009;33(2):103–10.

32. Mura T, Baramova M, Gabelle A, Artero S, Dartigues JF, Amieva H, et al. Predicting dementia using socio-demographic characteristics and the Free and Cued Selective Reminding Test in the general population. *Alzheimer's Res Ther*. 2017;9(1):1–11.
33. Jungwirth S, Zehetmayer S, Bauer P, Weissgram S, Tragl KH, Fischer P. Screening for Alzheimer's dementia at age 78 with short psychometric instruments. *Int Psychogeriatrics*. 2009;21(3):548–59.
34. Jacqmin-Gadda H, Blanche P, Chary E, Loubère L, Amieva H, Dartigues JF. Prognostic score for predicting risk of dementia over 10 years while accounting for competing risk of death. *Am J Epidemiol*. 2014;180(8):790–8.
35. Matthews FE, Chatfield M, Freeman C, McCracken C, Brayne C. Attrition and bias in the MRC cognitive function and ageing study: An epidemiological investigation. *BMC Public Health*. 2004;4:1–27.
36. Jokela M, Pulkki-Råback L, Elovainio M, Kivimäki M. Personality traits as risk factors for stroke and coronary heart disease mortality: pooled analysis of three cohort studies. *J Behav Med [Internet]*. 2014 Oct 8 [cited 2017 Aug 3];37(5):881–9. Available from: <http://link.springer.com/10.1007/s10865-013-9548-z>
37. Hagger-Johnson G, Roberts B, Boniface D, Sabia S, Batty GD, Elbaz A, et al. Neuroticism and cardiovascular disease mortality. *Psychosom Med [Internet]*. 2012 [cited 2017 Aug 3];74(6):596–603. Available from: <http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00006842-201207000-00007>
38. Phillips AC, Batty GD, Weiss A, Deary I, Gale CR, Thomas GN, et al. Neuroticism, cognitive ability, and the metabolic syndrome: The Vietnam Experience Study. *J Psychosom Res [Internet]*. 2010 Aug [cited 2017 Aug 3];69(2):193–201. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0022399910000218>
39. Deary IJ, Weiss A, Batty GD. Intelligence and personality as predictors of illness and death: How researchers in differential psychology and chronic disease epidemiology are collaborating to understand and address health inequalities. *Psychol Sci Public Interes*. 2010;11(2):53–79.
40. Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. *Bioinformatics*. 2014;31(9):1466–8.
41. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 2013;45(12):1452–8.
42. Kroenke K, Spitzer RL, Williams JBW, Lowe B. An Ultra-Brief Screening Scale for Anxiety and Depression: The PHQ-4. *Psychosomatics [Internet]*. 2009;50(6):613–21. Available from: <http://psy.psychiatryonline.org/cgi/doi/10.1176/appi.psy.50.6.613>

43. Eastwood S V., Mathur R, Atkinson M, Brophy S, Sudlow C, Flaig R, et al. Algorithms for the capture and adjudication of prevalent and incident diabetes in UK Biobank. PLoS One [Internet]. 2016;11(9). Available from: <http://dx.doi.org/10.1371/journal.pone.0162388>
44. UK Biobank. Anthropometry (Version 1.0) [Internet]. 2014. Available from: <http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=146620>
45. World Health Organization. Information sheet: global recommendations on physical activity for health 18 - 64 years old. 2011.

Table 1

Cognitive tasks during the brief test battery administered at UK Biobank baseline assessment centres

Cognitive test	Description	Scoring	n
Reaction time	A 'snap'-like card game in which participants were required to press a button as quickly as possible if the two cards appearing on the screen were identical.	Mean time (milliseconds) to correctly identify four matched pairs as part of 10 trials	496 787
Visual memory	The test required participants to identify six pairs of playing cards each with a different solid-black object or symbol. Individual cards were randomly assigned to positions on a 3x4 matrix and presented to the participant for five seconds, before being electronically 'flipped'.	Number of errors made in identifying correct pairs	497 986
Verbal-numerical reasoning	Referred to by UK Biobank as the 'fluid intelligence test', this was presented by 13 multiple-choice items (six verbal, seven numerical).	Scored from 0 to 13 (scores were reversed to indicate number of errors)	165 486
Prospective memory	Participants were informed that, at the end of the test battery, they would see four shapes of different colours on the screen and would be asked to touch a blue square, but instead they should touch an orange circle.	'1' for correct first time; '0' for incorrect first time	171 579

Note. The prospective memory and verbal-reasoning tasks were completed by approximately one third of the sample due to their later introduction to the UK Biobank baseline assessment.

Table 2

Constitutional and modifiable risk factors measured at UK Biobank baseline and used as covariates in the present study

Measure	Description	Reference group
Constitutional		
Educational attainment	<u>Touchscreen</u> : indicated by no education beyond secondary school ('1'), or having a college/university degree or equivalent ('0'), and considered in the present study to be a non-modifiable marker of premorbid intellectual ability	college degree
Family history of dementia	<u>Touchscreen</u> : self-reported father/mother/sibling with dementia diagnosis	no family history
Neuroticism	<u>Touchscreen</u> : scores from 12-item Eysenck Personality Questionnaire (EPQ-R)-Short Form. This relatively stable personality trait, viewed as a non-modifiable risk factor in the present study, was included due to its association with several vascular risk factors (36–38), independent of cognitive ability (39).	lowest quartile
APOE e4	<u>Genotyped blood data</u> : APOE e4 carrier status ('1' for e4 carrier and '0' for e4 non-carrier), excluding e2-e4 haplotypes	non APOE e4 carrier
AD polygenic risk	<u>Genotyped blood data</u> : AD polygenic risk score was calculated using the software PRSice (40) according to individuals' genotyped blood data and the summary statistics from International Genomics of Alzheimer's Project (41). The score was calculated according to effect size estimates of multiple single nucleotide polymorphisms (SNPs) found to be significant ($p < 0.01$) in case-control studies, with the exclusion of APOE-relevant SNPs given that we separately accounted for APOE genotype (see Supplementary Methods A for description of genotyping and quality control methods, and, accompanying Table S2).	lowest quartile
Modifiable risk factors		
Depression	<u>Touchscreen</u> : I. <i>Baseline risk</i> : a combined score of >3 on the first two items of the Patient Health Questionnaire (PHQ-2) that assess the two core criteria of depression experienced in the previous two weeks (42) II. <i>History</i> : if participant ever visited a GP for depression or anxiety	non-depressed non-depressed
Midlife hypertension	<u>Touchscreen</u> : self-reported regular use of blood pressure (neither self-reported hypertension nor physical measures were related to dementia).	no hypertension
Hyperlipidaemia	<u>Touchscreen and nurse interview</u> : self-reported regular use of cholesterol lowering medication, or, at nurse interview reported high cholesterol or named a cholesterol-lowering medication being taken (Table S3 lists these medications).	no high cholesterol
Diabetes	<u>Touchscreen and nurse interview</u> : self-reported regular use of insulin, or at nurse interview, reported diagnosis or taking of a diabetes-related medication (see Table S4 for list of medications); consistent with a previously published algorithm (43)	no diabetes
Obesity	<u>Anthropometric measures</u> (44): I. <i>Total obesity</i> : body mass index (BMI) of ≥ 30 II. <i>Central obesity</i> : waist-to-hip ratio (WHR) ≥ 0.85 women; ≥ 0.90 men	no obesity no obesity
Low physical activity level	<u>Touchscreen</u> : categorised as low (0), moderate (1), or high (2), according to World Health Organisation recommendations (45): low = activity less than moderate; moderate activity = 2.5 hr/week moderate or 75 minutes/week vigorous, or, $>$ once per week, or equivalent; high activity (for additional health benefits) = 300 minutes/week moderate or 150 minutes/week vigorous, or equivalent. Four self-report items from the International Physical Activity Questionnaire–Short Form covering the frequency and duration of moderate or vigorous physical activity during an average week.	high activity
Smoker	<u>Touchscreen</u> : self-reported never (0), ex-smoker (1), or current smoker (2)	never smoked

Table 3

Odds ratios (and 95% confidence intervals) for a first diagnosis of dementia made at least 3 years after baseline, according to established risk factors

	All-cause dementia <i>N</i> = 385 336 (591 cases) [§]			Alzheimer's disease <i>N</i> = 384 937 (192 cases)			Vascular dementia <i>N</i> = 384 837 (92 cases)		
	No dementia at follow-up %	Dementia at follow-up %	OR (95% CI)	No AD at follow-up %	AD at follow-up %	OR (95% CI)	No VaD at follow-up %	VaD at follow-up %	OR (95% CI)
Male	45.1	55.5	1.39 (1.18, 1.63)*	45.1	48.4	1.04 (0.78, 1.38)	45.1	62.0	1.81 (1.19, 2.75)
Age at baseline	56.3 yrs	64.2 yrs	1.21 (1.19, 1.23)*	56.3 yrs	64.4 yrs	1.22 (1.18, 1.26)*	56.3 yrs	64.4 yrs	1.22 (1.16, 1.28)*
Education, <i>no degree</i>	65.1	78.7	1.48 (1.22, 1.81)*	65.1	77.6	1.35 (0.96, 1.90)	65.1	83.7	2.09 (1.20, 3.64)
Constitutional									
<i>APOE</i> <i>e4</i> carrier	30.8	56.7	2.99 (2.43, 3.68)*	30.8	65.1	4.28 (2.97, 6.18)*	30.8	49.2	2.22 (1.33, 3.70)
AD PGS _{noAPOE} , <i>Q4</i> [†]	25.0	30.8	1.41 (1.06, 1.88)	25.0	38.1	2.11 (1.28, 3.47)	25.0	28.8	2.49 (1.03, 6.01)
Family history of dementia	12.3	23.2	1.65 (1.36, 2.00)*	12.3	29.7	2.29 (1.68, 3.12)*	12.3	28.3	2.16 (1.37, 3.41)*
Neuroticism, <i>Q4</i> [†]	24.6	24.5	1.45 (1.15, 1.84)*	24.6	22.9	1.13 (0.75, 1.69)	24.6	29.4	1.76 (1.01, 3.07)
Modifiable									
Depression									
PHQ-2	5.5	8.1	2.25 (1.67, 3.03)*	5.5	8.3	2.31 (1.38, 3.86)*	5.5	10.9	3.19 (1.65, 6.17)*
GP visit	33.8	38.2	1.47 (1.24, 1.74)*	33.8	37.0	1.34 (0.99, 1.80)	33.8	40.2	1.68 (1.10, 2.57)
Hypertension	19.8	38.9	1.39 (1.17, 1.64)*	19.8	32.3	1.04 (0.76, 1.41)	19.8	50.0	2.12 (1.40, 3.22)*
Hypercholesterolemia	16.9	35.7	1.36 (1.14, 1.61)*	16.9	28.1	0.97 (0.70, 1.34)	16.9	50.0	2.38 (1.56, 3.64)*
Diabetes	4.6	11.5	1.75 (1.36, 2.26)*	4.6	8.9	1.35 (0.82, 2.24)	4.6	18.5	2.93 (1.72, 4.99)*
Obesity [†]									
Overall	24.0	24.5	1.01 (0.84, 1.22)	24.0	22.4	0.90 (0.64, 1.26)	24.0	33.7	1.57 (1.02, 2.42)
Central	48.1	61.3	1.16 (0.97, 1.40)	48.1	53.7	0.91 (0.67, 1.25)	48.1	67.4	1.40 (0.87, 2.25)
Low physical activity	16.8	16.8	1.23 (0.98, 1.54)	16.8	12.5	0.82 (0.52, 1.27)	16.8	18.5	1.62 (0.91, 2.88)
Current smoker	9.9	11.2	1.53 (1.17, 2.01)*	9.9	8.9	1.13 (0.67, 1.90)	9.9	13.0	1.81 (0.95, 3.47)

Note. All OR (odds ratios) are from models that adjust for age and sex; those including *APOE* or AD PGS additionally adjust for genetic batch, array, assessment centre, and the first ten principal components to explain population structure. *Statistically significant after Bonferroni correction: 0.05/26 tests per outcome in univariate and multivariate regression models = *p*-value of 0.0019.

[§]All except for *APOE* *e4* and AD PGS_{noAPOE} (all-cause: n = 224 054, cases = 367; AD: n = 223 813, cases = 126; VaD: n = 223 746, cases = 59). [†]Q4 = highest quartile group; reference group is lowest quartile. ^{*}Overall obesity: BMI \geq 30; central obesity: WHR \geq 0.85 in women and \geq 0.90 in men.

Table 4

Prediction of incident dementia according to UK Biobank baseline cognitive tests, adjusted for age, sex and education

Cognitive test	Total dementia	Alzheimer's disease	Vascular dementia
<i>n (cases)</i>	<i>385 336 (591)</i>	<i>384 937 (192)</i>	<i>384 837 (92)</i>
Reaction time (1 SD milliseconds)	1.31 (1.22, 1.41)*	1.24 (1.09, 1.41)*	1.56 (1.32, 1.84)*
Visual memory (1 SD errors)	1.27 (1.18, 1.36)*	1.24 (1.10, 1.41)*	1.20 (1.00, 1.44)
<i>n (cases)</i>	<i>130 212 (115)</i>	<i>130 139 (42)</i>	<i>130 118 (21)</i>
Verbal-numerical (1 SD incorrect)	1.68 (1.36, 2.07)*	1.84 (1.30, 2.61)*	1.56 (0.97, 2.52)
Prospective memory (incorrect first time)	3.28 (2.26, 4.75)*	4.72 (2.51, 8.88)*	2.86 (1.20, 6.80)

Note. Odds ratios (and 95% confidence intervals) for incident dementia risk are association with a one standard deviation (SD) higher score on the cognitive measure (i.e. higher reaction time, or higher number of incorrect responses), except for prospective memory that is in association with an incorrect first response. *Statistically significant after Bonferroni correction: 0.05/26 tests per outcome in univariate and multivariate regression models = p -value of 0.0019.

Table 5

Incident all-cause dementia risk predicted by baseline cognitive tests, adjusted for constitutional and modifiable risk factors

	Model 1	Model 2	Model 3
<i>N</i> cases 397 485 646			
Reaction time (1 SD milliseconds)	1.31 (1.22, 1.40)*	1.30 (1.22, 1.39)*	1.28 (1.20, 1.37)*
include cases with APOE e4 status [§]	1.31 (1.20, 1.43)*	1.33 (1.22, 1.45)*	1.31 (1.20, 1.43)*
Visual memory (1 SD errors)	1.22 (1.14, 1.30)*	1.21 (1.13, 1.30)*	1.21 (1.13, 1.30)*
include cases with APOE e4 status [§]	1.21 (1.11, 1.32)*	1.19 (1.09, 1.30)*	1.19 (1.09, 1.30)*
<i>N</i> cases 134 623 122			
Verbal-numerical (1 SD incorrect)	1.69 (1.38, 2.07)*	1.71 (1.40, 2.10)*	1.67 (1.36, 2.04)*
include cases with APOE e4 status [†]	1.87 (1.42, 2.48)*	1.80 (1.10, 2.92)	1.72 (1.06, 2.81)
Prospective memory (incorrect)	3.23 (2.26, 4.64)*	3.29 (2.29, 4.71)*	3.16 (2.20, 4.55)*
include cases with APOE e4 status [†]	3.50 (2.15, 5.67)*	3.12 (1.35, 7.19)	3.00 (1.30, 6.96)

Note. Odds ratios (and 95% confidence intervals) for incident dementia risk are in association with a one standard deviation (SD) higher score on the cognitive measure, except for prospective memory that is in association with a correct first response. Model 1 covariates include age, sex, and education; Model 2 covariates include family history of dementia and neuroticism, and APOE e4 where stated (as in Tables S8-11) + Model 1 covariates; Model 3 adds modifiable risk factors (as in Tables S8-11) to Model 2, including current and past depression, hypertension, hypercholesterolemia, diabetes, and, smoking. *Statistically significant after Bonferroni correction: 0.05/26 tests per outcome in univariate and multivariate regression models = *p*-value of 0.0019). §Models include *n* = 230 539 (393 dementia cases). †Models include *n* = 76 326 (68 dementia cases).

Figure 1

ROC curves for the prediction of incident dementia according to: (1) age, sex and education, (2) + family history with *APOE e4* (left-hand graph) or without (right-hand graph) *APOE e4*; (3) + self-reported depression rating; (4) + cognitive test performance. Sample sizes are: (left-hand graph) 81 823 including 78 dementia cases; (right-hand graph) 145 068 participants including 141 dementia cases.